Communication evaluation in international organisations: methodology, influence and use

Glenn O'Neil

A thesis submitted to the Department of Methodology of the London School of Economics and Political Science for the degree of Doctor of Philosophy
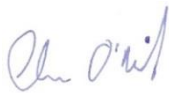
London, September 2015

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it). I consider the work submitted to be a complete thesis fit for examination. I authorise that, if a degree is awarded, an electronic copy of my thesis will be deposited in LSE Theses Online held by the British Library of Political and Economic Science and that, except as provided for in regulation 41 it will be made available for public reference.

I authorise the School to supply a copy of the abstract of my thesis for inclusion in any published list of theses offered for higher degrees in British universities or in any supplement thereto, or for consultation in any central file of abstracts of such theses. The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgement is made. This thesis may not be reproduced without my prior written consent. I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party. I declare that my thesis consists of 78,405 words.

I confirm that article 3 was jointly co-authored with Professor Martin W. Bauer and I contributed 80% of this work.
Signed:

Glenn O'Neil

As the candidate's supervisor I hereby confirm the extent of the candidate's contribution to the joint-authored papers as indicated above and in the Introduction below.
Signed:

Professor Martin W. Bauer

Statement of use of third party for editorial help: I confirm that my thesis was copy edited for conventions of language, spelling and grammar by Dr Sharon McClenaghan.

Acknowledgments

Firstly, I would like to thank my supervisor, Professor Martin W. Bauer for his excellent guidance and persistent dedication to what became for him an unintentional long term commitment of coaching and support. I would also like to thank current and past members of the Methodology Department for their support, comments and reviews of my thesis at its various stages, notably Dr Flora Cornish, Professor George Gaskell, Dr Dominik Hangartner, Professor Jon Jackson, Dr Jouni Kuha and Dr Jen Tarr.

I am grateful to colleagues and friends for their support, suggestions and comments on my thesis as it progressed: Ms Lois Austin, Ms Rosita Ericsson, Ms Patricia Goldschmid, Dr Sarah Grosso, Dr Jenevieve Mannell, Dr Sharon McClenaghan and Professor Tom Watson. I am also indebted to Ms Mohini Ghai Kramer of the International Committee of the Red Cross and Mr Victor Fernandez of the Office of High Commissioner for Human Rights for their support and collaboration in the cases carried out for this thesis. I thank also the International University in Geneva for their financial support for this thesis.

I have appreciated the long distance encouragement of my brothers and parents in Australia. Finally, I dedicate this thesis to my wife Roberta and our three children, Dante, Livio and Flavia; it was with their support, love and patience that I was able to complete this long but fruitful journey.

Abstract

International organisations (IOs) wield considerable influence in today's world. Distinguishing them from other actors are the new ideas they produce and communicate that can reframe global debates.   However, there is little research about how these organisations evaluate their communication activities. This thesis sets out to fill this gap by providing the first in-depth study of communication evaluation within IOs. The central question of this thesis is to assess the extent to which communication evaluation is possible within IOs with three specific questions: 1) the appropriateness and feasibility of communication evaluation methodology for IOs; 2) the influence of internal and external factors; and 3) the use of communication evaluation findings in IOs.

These questions are addressed in three distinct but interlinked empirical studies framed by a conceptual framework. Article 1 provides a 15 year review of communication evaluation within IOs through analysing systematically evaluation reports. Article 2 provides an analysis and reflection on the evaluation by this author of two communication campaigns of IOs. Article 3 provides a comprehensive study on the use of evaluation findings of these two communication campaigns.

This thesis found that a process of conceptualisation is needed to match the given communication activities to an appropriate evaluation methodology, implemented with a pragmatic, adaptive and participatory approach rather than imposing a standard set of methods. Evaluation was found only in a minority of IOs and there are challenges in matching evaluation methods to the range and complexity of their communication activities.

Internal factors of organisational context and communication goals were found to be important and to equally impede and enable the evaluation process. Improving the efficiency of communication activities was the main use seen of evaluation findings, occurring in unexpected and often opportunistic ways, but nevertheless an indication of a specific contribution of evaluation for communication professionals.

Table of Contents

List of Figures and Tables

Chapter 7

Annexes

1. Introduction

1.1. Background

Communication is increasingly recognised as a powerful and strategic tool for organisations, companies and individuals to influence, develop relationships and promote their values and causes (Holtzhausen & Zerfass, 2014; Sriramesh, 2009). In today's globalised environment, international organisations (IOs) have become adept at using communication; the World Health Organisation's (WHO) campaigns to change health practices; the ability of World Wildlife Fund (WWF) and other environmental organisations to put "green" issues high on the public agenda; and the attention and pressure on human rights abuses by governments that Amnesty International and other organisations can bring through their communication actions are just some examples (Keck and Sikkink, 1998; Kickbusch, 2003). It has been argued that what distinguishes IOs from military, political and economic actors is one key attribute: ideas. "Human rights", "climate change", "global justice" and "development goals" are prominent examples of theoretical concepts that IOs have effectively communicated to reframe global debates (Keck and Sikkink, 1998; Willetts, 2010).

However, little attention has been paid as to how IOs manage and evaluate their communication activities (Lehmann, 1999; Manheim, 2011; Schwarz & Fritsch, 2014; Sireau, 2009; Tkalac & Pavicic, 2009). Academia and the communication industry have produced theories, models and guides on evaluating communication activities that have yet to be adopted widely (Macnamara, 2014; Cutlip, Center & Broom, 2006). Together with the extensive research, debates and guidance of the general evaluation field, this represents a vast domain of knowledge and practices whose relevance and applicability for communication evaluation within IOs is little understood.

This thesis sets out to fill this gap by providing the first in-depth study of communication evaluation within IOs. The idea that underpins this thesis is that a study of methods and their use is not sufficient to understand communication evaluation within IOs. More so,

the phenomenon that is evaluation has to be studied in its totality, in relation to the environment in which it is carried out and the various factors that can influence its success or not.   In this regard, different theories and research methods need to be considered, as is the case throughout this thesis.

This author's background brings a particular perspective to this research. As an independent evaluator for over ten years with a specialisation in communication and the non-profit sector, this author has carried out some 100 evaluations covering over 50 countries and has constantly been struck by the gap between the theory and practice of evaluation; the linear and logical frameworks imposed and the complex, varying and often messy contexts in which they are intended to be applied.

1.2. Research questions

In mentioning this perspective, the point of this author is not to persuade the reader as to a given position or orientation but more so to set the scene for this thesis that will aim to provide both a contribution to the theory and practice of communication evaluation. This will describe first-hand the possibilities and limitations of where theory and practice meet through empirical research which forms the core of this thesis. Thus, the central question of this PhD study is as follows:

- To what extent is communication evaluation in IOs possible?

This central question is developed into several specific questions:

- What evaluation principles, methods and procedures are appropriate and feasible for IOs to evaluate their global communication programmes and campaigns?

- To what extent do contextual, organisational and human factors influence the ability of IOs to evaluate their global communication programmes and campaigns?

- How are communication evaluation findings used within IOs and what factors enable and impede their use?

In answering the central and specific questions, the PhD study will make an important contribution to a current gap at the intersection of three bodies of knowledge where no known empirical research has been carried out, as illustrated in the following diagram:
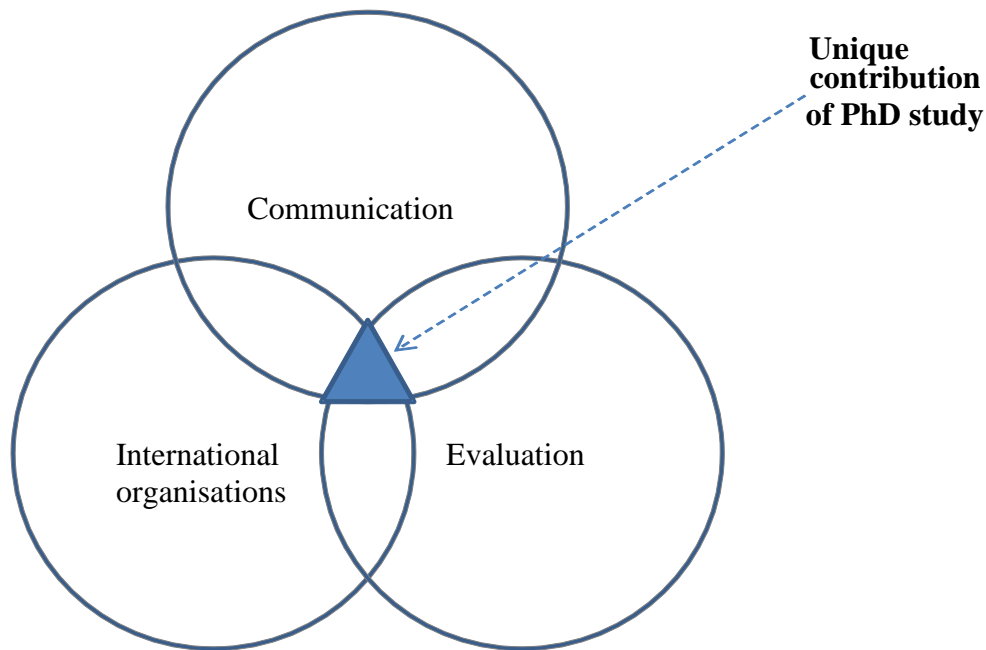
Figure I: Unique contribution of the PhD study

The contribution of this thesis to these three bodies of knowledge is further detailed in chapter 3 below.

1.3. Structure of the thesis

This thesis is in line with the Methodology Department's guidelines for a paper-based PhD thesis. This thesis consists of an introductory chapter (chapter 1), a literature review chapter (chapter 2), a chapter which describes the contribution of the thesis (chapter 3), three articles of empirical research (chapters 4, 5 & 6) and linking texts between them (interludes 1, 2, 3 & 4), and a conclusion chapter (chapter 7).

The substantive component of this thesis is the three articles. A shorter version of article 1 (chapter 4) has been published in the *Public Relations Review* (O'Neil, 2013) a peer-

reviewed journal and the remaining two articles (chapters 5 & 6) are currently being submitted for publication (O'Neil, 2015; O'Neil & Bauer, 2015).

As each article is written as a stand-alone paper, there is some repetition, particularly of major theoretical discussions and the literature. The articles also reference each other.

A description is provided below of each of the three articles that operationalise the central and three specific questions in addition to the conclusion chapter:

Article 1: 15 year review of communication evaluation within international organisations

The purpose of this article is to understand how IOs have evaluated their communication activities and to what extent they have adhered to principles of evaluation methodology over a 15 year period (1995-2010). This is carried out through a review of available evaluation reports and guidelines. The steps and protocols of the systematic review methodology guide this article.

Article 2: Implementing communication evaluation methodologies for two international organisations

The aim of this article is to consider what evaluation methodology is appropriate and feasible for communication activities of IOs. This is done through the experience of this author in evaluating two international communication campaigns of the Office of High Commissioner for Human Rights (OHCHR) and the International Committee of the Red Cross (ICRC), considering the internal and external factors that influence implementation. The author's own experiences are reflected upon through an "insider-outsider" research approach and a mixed methods strategy is used for the evaluation of the two campaigns documented in this article.

Article 3: The use of communication evaluation findings in two international organisations

Understanding the usage of evaluation findings is the focus of this article. The author returns to the OHCHR and the ICRC to examine how the campaign evaluation findings have been used some four years later. Instances of use are documented and verified

using a qualitative methodology centred on a typology of intended/unintended use, level, attributes of change and influences.

Conclusion chapter

The conclusion chapter covers six main points. Firstly, it reviews how the evidence presented in the articles responds to the three research questions. Secondly, it considers the theoretical implications and thirdly, the methodological lessons. Fourthly, recommendations are provided for the practice of communication evaluation within IOs. Finally, directions for further research will be suggested and broader implications outlined. The operationalisation and linkages between the research questions and the articles and conclusion are illustrated in Figure II.
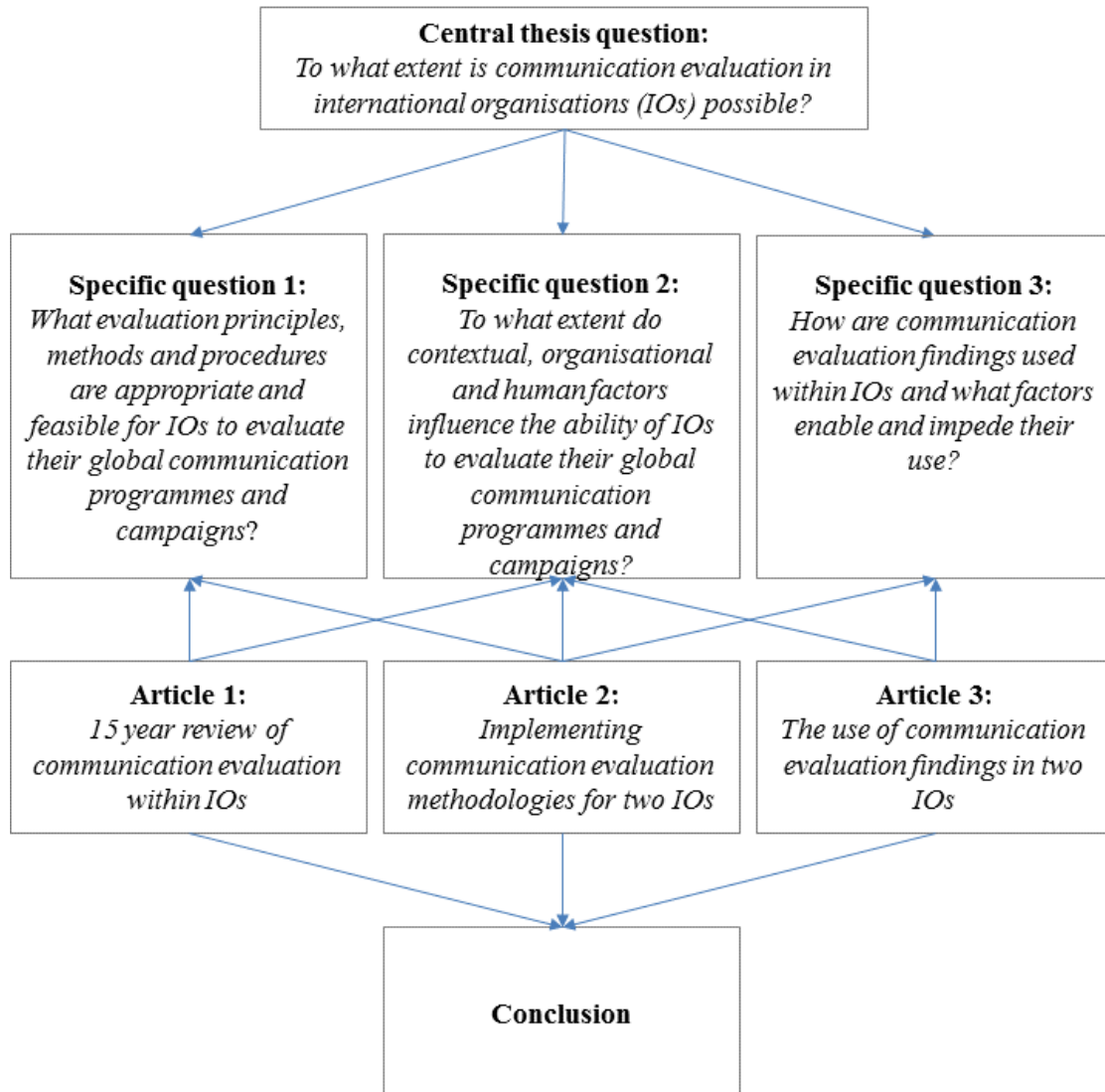
Figure II: Operationalisation of research questions

1.4. Theoretical framework of the thesis

This is a study on evaluation and provides a contribution to both the theory and practice of communication evaluation and the broader fields of organisational communications, international relations and evaluation. The main theoretical frame of reference for this study is systems theory and its application to organisations. The study commences with a review of literature that defines some of the relevant key concepts of communication, evaluation, communication evaluation and the organisations under study, IOs. The current state of this field and pertinent issues are explored and discussed. At this stage a conceptual framework for communication evaluation of IOs is proposed. The framework, guiding the research, is composed of four components of communication evaluation linked to internal and external factors, explained and developed further in chapter 3.

The first component (Methodology) of the framework is considered in article 1, where through a review of existing evaluation reports and guidelines, an understanding is sought of current evaluation practices and how they comply with broad principles of evaluation methodology.

The second component (Implementation) is considered in article 2 where the experience of this author in carrying out two evaluations is reflected upon considering the limitations of methodologies, design and methods in reality and the influences on evaluation.

The third and fourth components (Findings and Use) are considered in the third article through investigating the use of evaluation findings in two organisations (article 3).

The conclusions reconsider the conceptual framework drawing from the findings of the three articles with the aim of proposing methodological learnings and recommendations for the practice of communication evaluation of IOs. In totality, this provides both a theoretical basis and operational learnings for linking the framework to practical application for communication evaluation.

1.5. Methodological approach of the thesis

The articles that make up this study utilise a variety of methodological approaches to address the key questions. Article 1 is guided by a systematic review methodology, a type of literature review to identify and synthesize evidence, in this case to assess the methodological quality of existing evaluations. Article 2 draws upon the "insider-outsider" research approach to report on the multiple types of qualitative and quantitative tools used to evaluate two communication campaigns by this author. Article 3 uses qualitative methods to investigate in-depth instances of evaluation use that are categorised and analysed based on typology of intended/unintended use created from multiple sources and the author's own conceptual framework.

Of note, all articles produce original empirical research with methods and data that have not been widely used together in this field. The research and data was generated by this author with the exception of the following: another student provided support for the double-coding process of article 1; some research results of article 2 were generated by the staff of the organisations under study; and article 3 was co-authored with Professor Martin W. Bauer. For the latter, this author contributed some 80% of the article and the remaining 20% was the contribution of Professor Bauer.

This study is part of the communication studies discipline, notably the sub-discipline of organisational communications although it draws from other disciplines including international relations, organisational behaviour, social psychology, systems theory and evaluation itself (considered a "trans-discipline" or "almost" a discipline) (Clarke, 1999).

1.6. References

Clarke, A. (1999). *Evaluation research: An introduction to principles, methods and practice*. London: Sage Publications.

Cutlip, S. M., Center, A. H., & Broom, G. M. (2006). *Effective public relations* (9th ed.). Upper Saddle River, NJ: Prentice Hall

Holtzhausen, D., & Zerfass, A. (Eds.). (2014). *The Routledge handbook of strategic communication*. Chicago, IL: Routledge.

Keck, M., & Sikkink, K. (1998). *Activists Beyond Borders: Advocacy Networks in international politics*. Ithaca, NY: Cornell.

Kickbusch, I. (2003).The Contribution of the World Health Organization to a New Public Health and Health Promotion. *American Journal of Public Health*, Vol. 93, No. 3: 383-388.

Lehmann, I. (1999). *Peacekeeping and Public Information, Caught in the crossfire*. London: Frank Cass.

Macnamara, J. (2014). The 'toe bone to the head bone' logic model to connect public relations and corporate communication to organisation outcomes. *Prism* (online PR Journal), Vol. 11, Issue 1. Retrieved from: http://www.prismjournal.org/fileadmin/11_1/Macnamara_2014.pdf

Manheim, J. (2011). *Strategy in information and influence campaigns: how policy advocates, social movements, insurgent groups, corporations, governments, and others get what they want*. New York, NY: Routledge.

O'Neil, G. (2013). Evaluation of international and non-governmental organizations' communication activities: A 15 year systematic review. *Public Relations Review,* 39, 572-574

O'Neil, G. (2015). Implementing communication evaluation methodologies for two international organisations. *Submitted for publication.*

O'Neil, G., & Bauer, M.W. (2015). The use of communication evaluation findings in two international organisations. *Submitted for publication.*

Schwarz, A., & Fritsch, A. (2014). Communicating on Behalf of Global Civil Society: Management and Coordination of Public Relations in International Nongovernmental Organizations. *Journal of Public Relations Research*, 26(2), 161-183.

Sireau, N. (2009). *Make Poverty History: Political Communication in Action*. Hampshire: Palgrave Macmillan.

Sriramesh, K. (2009). Globalisation and public relations: The past, present, and the future. *Prism,* 6(2). Retrieved from:
http://www.prismjournal.org/fileadmin/Praxis/Files/globalPR/SRIRAMESH.pdf

Tkalac, A., & Pavicic, J. (2009). Nongovernmental organizations and international public relations. In K. Sriramesh & D. Verčič (Eds.), *The global public relations handbook: Theory, research, and practice.* 2nd ed. (pp. 807-821) New York, NY: Routledge.

Willetts, P. (2010). *Non-governmental organizations in world politics: the construction of global governance*. London: Routledge.

2. Literature review

A review of the current literature has been carried out which commences with a definition of the main concepts and terms, followed by an overview drawn from the relevant academic and grey literature. The literature review draws from academic sources as well as grey literature, namely guidelines, studies and reports of the evaluation and communication evaluation fields, to supplement the lack of research in the former.

2.1. Concepts and definitions

The population of interest to this study are IOs that make up what is referred to as the "international public" or "international non-profit" sector (Charnovitz, 2006; Reinalda, 2013). In the international relations field, this sector can be interpreted broadly and include all non-state actors, such as international non-governmental organisations (INGOs) and intergovernmental organisations (IGOs) but also rebel movements, non-recognised authorities, transnational corporations, international criminal networks and even terrorist groups (Dijkzeul & Beigbeder, 2003; Martens, 2002). A more narrow interpretation would be to consider a subset, such as those INGOs playing a role in international affairs; or only United Nations (UN) organisations; or only IGOs (Dijkzeul & Beigbeder, 2003; Simmons, 1998). For this study, this author uses a middle-way definition of IOs found in several key international references (Dijkzeul & Beigbeder, 2003; Reinalda, 2013; Union of International Associations, 2014), namely, the major INGOs and IGOs active at the global level. This population includes three groups: INGOs with category 1 (general consultative status) with the UN; IGOs that have a standing invitation to participate as observers in the sessions and the work of the UN General Assembly; and the recognised entities of the UN system (also IGOs by definition). As of January 2011, 230 organisations corresponded to these criteria (a complete list is found at annex 1 of this thesis).

IGOs are created by governments to undertake a variety of functions including cooperation, monitoring, dispute settlement or humanitarian intervention (Mingst, 2004).

The World Bank is an IGO, as is the North Atlantic Treaty Organisation (NATO), the European Union (EU) and the United Nations Educational, Scientific and Cultural Organisation (UNESCO). INGOs are private organisations, normally with a voluntary aspect, whose members come together for a common purpose and are active internationally (Mingst, 2004; Simmons, 1998). INGOs undertake a variety of functions including acting as advocates for specific policies, mobilising publics, monitoring or providing humanitarian assistance. Greenpeace is an INGO as is Doctors without Borders (*Médecins Sans Frontières*) and so is Amnesty International. A key distinction between IGOs and INGOs is that the mandate of an IGO is based on a formal agreement between governments whereas an INGO normally has no direct mandate from governments (McLean, 2000).

It has been put forward that the distinction between INGOs and IGOs is increasingly blurred and far too much emphasis in the literature has been placed on the differences rather than the similarities between these organizations (Reinalda, 2013; Willetts, 2010). The key distinction between INGOs and IGOs, that is, the formal mandate given to IGOs implying that they are not independent of the states that created them, has been challenged by studies of how IGOs actually behave; "many I[G]Os exercise power autonomously in ways unintended and unanticipated by states at their creation" (p. 699, Barnett & Finnemore, 1999). This has been a longstanding discussion, with Reinalda (2013) citing a 1964 study by Haas of the International Labour Organization (ILO) that concluded that it had acquired independence from the states that created it and was effectively intruding into the national domains of states. More recent studies provide examples of IGOs such as the United Nation High Commissioner for Refugees and the International Committee of the Red Cross that found them in direct confrontation with states (Barnett & Finnemore, 1999; Willetts, 2010). Further, the definition of an IO as either an INGO or an IGO is not always categorical with the example cited of the World Conservation Union, which shifted from being treated as an INGO to an IGO due to its changing status with the UN (Willetts, 2010). According to Willetts, this necessitates the creation of a third type of IO, what he labels as a "hybrid international organisation" (Willetts, 2010, p. 73).

The similarities of INGOs and IGOs have led some scholars to study them together under the label of "international organisations" as does this study. Some of the key similarities between these organisations include: they provide goods and services which serve an international public purpose; they define global concepts, tasks and interests; they address comparable or the same societal issues; and importantly for this study, they function as forums for advocacy, communication and information exchange (Barnett & Finnemore, 1999; Dijkzeul & Beigbeder, 2003).

Within this population of interest, communication activities are the relevant functional area of concern for this study. These activities do not refer to all aspects of an organisation's communications internally and externally, but more so to the programmes, projects, actions and campaigns that are part of the management of communications between an organisation and its publics (Grunig, 1992).

Alternative terms are often used to describe these activities: Communication management, corporate communications, public relations, public information or public affairs. The communication function within an organisation normally includes specific sub-functions to carry out these activities such as media relations, public affairs, publicity, marketing support, online communications, identity or reputation management and media production. Within IOs, there may be additional sub-functions related to their specific nature such as public awareness/information and education functions (Lehmann, 1999). Further, as IOs are working globally and across multiple countries and culture, their communication falls under the definition of international public relations or communication (Banks, 1995; Culbertson, 1996). IOs have increasingly used communication to profile themselves, influence issues and build relations, as discussed in the next section (Barnett & Finnemore, 1999; McLean, 2000; Mingst, 2004; Schwarz & Fritsch, 2014; Welch, 2001).

Evaluation is considered by this study as "The systematic assessment of the operation and/or the outcomes of a program or policy" (Weiss, 1998, p. 4). Organisational theory considers evaluation as an "institution", a frame of action within organisations, a

regulative element that has developed its own "field", a community of organisations that has a common meaning system (Højlund, 2014a; Scott, 1995). Evaluation in this form is a relatively new phenomenon in society. There is general agreement that the beginning of modern evaluation was in the 1960s with the organised appraisals of the large social programmes in the United States (Pawson & Tilley, 1997; Weiss, 1998). In the next decades, evaluation spread to many other sectors and globally, including the non-profit sector and its communication activities (Shadish, Cook & Leviton, 1991). In the past fifty years, the evaluation field has grown dramatically, accompanied by an industry of consultancies, consultants, conferences, specialised media and professional associations (Stern, 2006).

Conceptually, evaluation is generally accepted as having its roots in applied social research that has now developed into a "trans-discipline" or "almost" a discipline (Clarke, 1999). Although utilising methods adapted from social sciences, it is seen as distinguishing itself by the purpose for which these methods are used, for example, accountability and social inquiry (Alkin, 2012).

In the communication field, evaluation has also been defined by the US-based Institute of Public Relations (IPR) as:

> A form of research that determines the relative effectiveness of a public relations campaign or program by measuring program outcomes (changes in the levels of awareness, understanding, attitudes, opinions, and/or behaviours of a targeted audience or public) against a predetermined set of objectives that initially established the level or degree of change desired. (Stacks, 2006, p. 7)

The IPR definition speaks of evaluation as measuring "program outcomes" but does not mention program processes, outputs or impact that are considered as alternative focuses of evaluation in communication (Communications Consortium Media Centre[CCMC], 2004). This narrower definition has been criticised as it potentially excludes the involvement of communication professionals whose main interest would be in these

alternative focuses to improve their activities and also because it favours evaluation as a distinct action carried out by external research professionals at the end of an activity (Watson and Noble, 2007).

2.2. The influential role of IOs and their communications

Despite numbering in several hundreds, IOs wield considerable influence in today's world: their ability to persuade major governments to take significant foreign policy decisions (Busby, 2007); their capacity to mobilise global publics to react to human rights abuses (DeMars, 2005); the influence of IOs such as the EU and the World Health Organisation (WHO) on the delivery of medical services and health practices of publics; and the key role of the International Monetary Fund and the World Bank in economic growth in developing countries are just several examples (Barnett & Finnemore, 1999; Duncan, 2002; Kickbusch, 2003). Although the veracity of such influence is questioned by realist theory in international relations (Waltz, 1979), constructivist theory counter that IOs are today some of the most important influences on states, equal or more influential than political, military and economic actors (Barnett & Finnemore, 1999; Elman and Elman, 2003).

An important difference with other actors is the new ideas and norms that these organisations produce and promulgate that consequently can reframe international, regional and national debates. The development of human rights is a case in point. In the past 60 years, human rights has been transformed from a set of theoretical concepts into a series of international treaties and mechanisms largely at the urging of IOs, with the result today, where individuals can bring their own governments before intergovernmental committees for human rights' violations (Willetts, 2010). This is not denying the role these organisations play in the technical assistance they provide governments and populations. However, if this role is set aside, a key activity for these organisations is the collection and analysis of information - and its communication (Keck and Sikkink, 1998).

However, the literature to date is largely silent on how these organisations manage and evaluate their communication activities (Lehmann, 1999; Manheim, 2011; Schwarz & Fritsch, 2014; Sireau, 2009; Tkalac & Pavicic, 2009). Studies to date of IOs have also been criticised for their consideration of these organisations as "self-contained units" (Dijkzeul & Beigbeder, 2003, p.7) with limited attention paid to interrelations and contextual aspects.

In this respect, the substantial contribution of this thesis is in providing an in-depth study into communication evaluation of IOs, and going beyond a "self-contained unit" approach by considering their settings and environment.

2.3. Globalisation of communication

The way organisations are communicating has changed massively in the last twenty years, regardless of whether they are public or private, national or international. Research largely points towards two inter-related developments: Globalisation and Information and Communications Technologies (ICTs) (Sriramesh, 2009). As organisations have expanded their reach across countries and regions, often into new markets with vast differences in political, social, economic and cultural environments, the implications for communication activities to adapt have been substantial (Rantanen, 2005; Sriramesh & Verčič, 2007; Tkalac, & Pavicic, 2009). The developments in ICTs that have accompanied globalisation has had notable impacts on communication including the ability of audiences to interact amongst themselves and with organisations, the broadening of information sources available, the ease of forming coalitions and the emergence of new ways of online social interaction, participation and mobilisation (Deibert, 2000; Movius, 2010; Wellman, 2002).

As a consequence, this has opened the potential for organisations to be more global, strategic and interactive in their communication – and in principle more evaluable - even if this is yet to be fully realised (Grunig, 2009; Sriramesh, 2009). At the same time, it also implies that organisations operating globally such as IOs have an inevitable

complexity in their communication and this can impact on their capacity to manage and evaluate.

In this study, the increasing complexity and globalisation of communication for IOs is considered in relation to its implications for evaluation; how does operating in complex environments and the global nature of the organisations impact on their ability to evaluate. Research is yet to explore fully how operating globally impacts on communication evaluation methodology and this study will add to the literature on this point.

2.4. Evaluation methodology, paradigms and management

This section describes five issues of current debate within the evaluation field of relevance to this study: understanding evaluation methodology; competing evaluation paradigms; orientations and incompatibilities within evaluation; the evaluation institution, system and policies; and the use of evaluation findings.

2.4.1. Understanding evaluation methodology

Within the literature, consensus is found around the notion that evaluation methodology is concerned with the principles, approach and procedures, including the methods used and steps taken. It is within this consensus that this study situates its understanding of evaluation methodology. Weiss (1998) viewed evaluation methodology as encompassing the design, measurement and analysis aspects of evaluation; Scriven (1991) as principles for general investigation, analysis and practical procedures; and Davidson (2005) as a set of principles and procedures that guides evaluation. There is also interest to identify harmonising elements to advance what is known as the "global evaluation theory" (Shadish, Cook, & Leviton, 1991), which would then serve as a common basis for procedures and principles for evaluators, although it remains underdeveloped (Rossi et al., 2007).

Principles and procedures to guide evaluation are numerous and have been developed at different levels and for different purposes. These include: ethical principles for evaluators, the practice of evaluation in different sectors, procedures for capacity building and

participation, the procedural steps and management of the evaluation process, guidance on designs, criteria, methods, analytical techniques and the use of evaluation findings (Greene, Mark & Shaw, 2006; Patton, 2008; Rossi et al., 2007; Shadish, Newman, Scheirer & Wye, 1995; Weiss, 1998).

Yet, the overall approach that guides evaluation and from which are derived these principles and procedures is diverse, fractured and competitive and has been labelled as the "paradigm wars" (Datta, 1994).

2.4.2. Competing evaluation paradigms

Originating from the 1970s, the "paradigm wars" pitted positivists against constructivist and was largely concerned with the evaluation designs adopted and the consequent methods used, replicating similar debates in the broader social sciences (Habermas, 1974). Over time, this debate was further nuanced with different paradigms and schools of thought emerging which crystallized around four main paradigms competing for attention: 1) Analytic-empirical-positivist-quantitative; 2) constructivist-hermeneutic-interpretive-qualitative; 3) critical theory-neo-Marxist-postmodern-praxis; and 4) eclectic-mixed methods-pragmatic (Reeves, 1997).

The dominant paradigm adopted by an evaluation, either knowingly or unknowingly, is seen as important as it influences the overall methodological approaches such as the design and methods used. It has been argued that there is validity in each of these paradigms and those undertaking evaluation usually adhere to a model that makes the most intellectual sense to them and that is also potentially adaptable. Further, it has been challenged that the dominant paradigm dictates the methods used; more so that methods are practical responses to solve specific problems and are relatively free of philosophical positions (Alkin, 2012; Maxwell, 2012).

In the absence of a common basis, tension has arisen again amongst academia and practitioners with the "paradigm wars" re-emerging as a debate on causality and evidence. On one side, there are those claiming that randomised-control trials are the "gold standard"

for determining causality. On the other side, there are those whose see such trials as having a very limited role and propose theory-based methods and approaches as alternatives (Picciotto, 2012).   A middle ground is also proposed by Patton (2011):

> The real gold standard is methodological appropriateness, namely, matching methods to the nature of the question and the purpose of the evaluation (p. 290).

This study provides a reflection on the position of communication evaluation within IOs in relation to the competing paradigms debate and its influence (or not) on methodological approaches taken.

2.4.3. Orientations and incompatibilities within evaluation

This ongoing debate has been accompanied by the emergence of different orientations or schools of thought that offer direction and guidance to those carrying out evaluation that are not mutually exclusive, such as realist evaluation (Pawson & Tilley, 1997), real-world evaluation (Bamberger, Rugh & Mabry, 2006), participatory evaluation (Cousins & Chouinard, 2012), developmental evaluation (Patton, 2011) and complexity evaluation (Rogers, 2008).

What a number of these orientations share in common is their recognition of the limits of evaluation theories and their application in reality to what Schwandt (2003) called the "rough ground of paradox and contingency, ambiguity and fragmentation" (p. 361).

 A key challenge identified is that evaluation is essentially a logical and linear process that does not necessarily match the complex activities that it is seeking to evaluate across multi-institutions (Rogers, 2008) and that methodology and consequent methods have been developed with little understanding of how they relate to context, organisational and human factors (Fitzpatrick, 2012; Mathison, 1994).   Drawing from systems theory, alternative approaches have been developed such as non-linear models and adapted evaluation designs (William & Hummelbrunner, 2010; Williams & Imam, 2007).

Nevertheless, Sanderson (2000) doubts whether evaluation has moved very far away from the positivist perspective given that in practice evaluation frameworks are still dominated by a linear positivist structure, as seen in the dominance of the logical framework. At the same time, Hall (2014) found that a given evaluation practice, such as the logical framework, can be aligned with different ideals and beliefs about evaluation and is not necessarily exclusive to one or the other. Reynolds (2015) found that evaluation is not yet able to assess activities from a perspective of being integrated or interrelated within the larger organisation. At the same time, Reynolds recognises that there has been a move in evaluation away from linear thinking towards considering the impact of interventions in their contexts representing "a shift from being systematic to being more systemic" (p. 71).

Concern has also been raised over the incompatibility between the increasing adoption of linear driven results-based management systems by organisations (Mayne, 2007) at the same time as a push for more participative, flexible, context-sensitive evaluation approaches (Lennie & Tacchi, 2011). The evaluation process is pressured simultaneously from the "top-down" that demands standardized procedures and from the "bottom-up" of contextual factors and programme adaptation that exist in reality, placing evaluation in the "muddled middle" (p.177) as labelled by Patton (2011).

This study will add to the existing literature on these limits and incompatibilities by providing a further perspective of the "muddled middle" from a particular sector (IOs) for a specific function (communication).

2.4.4. The evaluation institution, system and policies

The place of evaluation within organisations, its institutionalization and policies has been labelled as one of the most important issues currently facing evaluation although not yet receiving the attention warranted (Bamberger, Rugh & Mabry, 2006; Trochim, 2009). It is put forward that evaluation has become an 'institution' within organisations, that the act of evaluation has taken the form of a 'ritual' that is necessary for political and administrative validity which is little questioned (Dahler-Larsen, 2012; Schwandt, 2009).

At the same time, in a study of evaluation within EU bodies, it was found that the institutionalisation of evaluation was largely a positive influence in making evaluation practice more systematic and routine, notably in the use of evaluation findings (Højlund, 2014b). The literature indicates that the evaluation institution is operationalised in organisations usually through an evaluation unit that sits at the centre of the larger evaluation system which spans across the organisational entities, promulgating permanent and systematic, formal and informal evaluation practices with the purpose of informing decision making and providing oversight, even if little is known about the system, its interrelations with other components and systems (Højlund, 2014b; Leeuw & Furubo, 2008).

Evaluation policies are seen as important for the institution and system as they guide how evaluation happens, such as which methods are preferred, the level of participation desired and how evaluation findings are used. Decisions are taken on which principles and procedures to adopt, and in doing so, which paradigms, approaches and orientations are favoured. Evaluation policies have been seen to be a mix of informal and formal, written and unwritten with limited attention paid to date as to how they are interpreted and implemented in practice (Christie & Fierro, 2012; Mark, Cooksy & Trochim, 2009; Trochim, 2009).

Research on decision-making for methodological choices has focused on the role of the external evaluator (engaged by the evaluation unit to carry out evaluation) and what influences their choices, such as issues of feasibility, legitimacy of the evaluation process and the perceived expectations of audiences, rather than on decision-making as a whole within the evaluation system (Fitzpatrick, Christie & Mark, 2009; Tourmen, 2009).

A key differentiation between this study and previous research into evaluation methodology and practice is its consideration of evaluation within the broader context of the evaluation institution, systems and policies in relation to a specific function (communication), rather than as an isolated process or practice, bringing an additional perspective to this field. Further, in considering methodological choices, this study will

go beyond the evaluator's perspective on what is feasible and appropriate for them; more so what influences these perspectives and of the system as a whole.

2.4.5. Use of evaluation findings

The use of evaluation findings has attracted considerable interest for research, linked to the notion put forward that evaluation should be judged by its utility and actual use (Patton, 2008). Theoretical frameworks studying evaluation use have been produced with the most common being a categorization of four types of use: instrumental, conceptual and symbolic use, with process use added more recently (Mark and Henry, 2003; Patton, 2008; Shulha & Cousins, 1997). In studies to date, there is an expectation that instrumental use will dominate, that is, evaluation will lead directly to changes and decision-making. However, usage has mainly been found to be conceptual, such as new learning drawn from evaluation results but no direct action occurs. Usage has also been found to be symbolic, such as justifying a pre-existing position or to be the basis for action or inaction (Mark and Henry, 2003; Mark and Henry, 2004; Shulha & Cousins, 1997).

The study of evaluation use also has a pragmatic application, in that scholars have endeavoured to determine what are the factors that are likely to predict greater use of evaluation results. A commonly used typology has been the set of 12 factors developed by Cousins and Leithwood (1986). Six factors concern the evaluation implementation: Evaluation quality, credibility, relevance, communication quality, findings and timeliness; and six factors focused on the decision or policy setting: Information needs, decision characteristics, political climate, competing information, personal characteristics and receptiveness/commitment. An additional category of "stakeholder engagement" was added to this classification by Johnson et al. (2009). Empirical reviews of evaluations on these factors have varied in their findings with the following highlighted as being more influential on use: credibility of the evaluators and their products; decision-making characteristics; policy setting; and stakeholder engagement (Cousins and Leithwood, 1986; Højlund, 2014b; Johnson et al., 2009).

A limitation of the research to date has been that its focus has virtually always been on the evaluation process and does not take into account sufficiently the organisational, context and human factors that are potentially stronger influences on use (Contandriopoulos & Brousselle, 2012; Højlund, 2014b). There has also been little understanding of the underlying processes of how use occurs, how evaluation findings are received and understood and its influence on change within organisations, with the exception of the framework produced by Henry and Mark (2003 & 2004) of mediators and pathways for change at three levels, individual, intrapersonal and collective (organisational). However, this framework is yet to be operationalised in empirical studies, with two known attempts not succeeding (Johnson et al., 2009; Weiss, Murphy-Graham & Birkeland, 2005).

When considering evaluation use, this study will investigate use in relation to these broader organisational, context and human factors, based on the notion, drawn from systems theory that an organisation is made up of interrelated parts and that processes, such as evaluation and its use cannot be studied in isolation within the organisation and from its environment. Further, article 3 will integrate and apply the framework of Henry and Mark (2003 & 2004) in its empirical study of evaluation use in IOs.

2.5. Specificities of communication evaluation

The above section considered the development and current debates of the evaluation field in general and those particularly relevant to this study. This section considers the development of evaluation particularly for the communication function within organisations with an emphasis on models and designs of communication evaluation, efforts to categorise communication effects, commonalities, prevalence and limitations of the field.

2.5.1. Development of communication evaluation

The development of evaluating communication activities has been well documented (Gregory & Watson, 2008; Watson & Noble, 2007). Interest in evaluating communication activities began in the 1920s, when the first systematic communication

activities were launched in the United States. For the first decades, evaluation focused on understanding the ability of communication activities to influence attitudes and public opinions, notably through social-psychological studies. As mass communication theory developed, so did evaluation, largely following the broad developments in theory; from the "hypodermic-needle effects" model of the direct, universal and all powerful influence of mass media to the "minimum effects" model that downplayed the influence of mass media and focused on the obstacles in influencing audiences (Salwen & Stacks, 1996).

As structured and organised communication activities were adopted by companies, governments and organisations, academics focused further on understanding how these new activities functioned and what did they achieve. Models and concepts of the different types of communication activities were formulated (Grunig & Hunt, 1984) and consequently evaluation models, methodology and tools, adapted largely from social sciences, were developed to match what were becoming generally acceptable concepts of what communication activities were supposed to "do" and "achieve" (Broom & Dozier, 1990; Cutlip, Center & Broom, 1994; Macnamara, 1992; Watson, 1997). This was also linked to the gradual move of communication professionals from "technicians" to "manager" and more valued placed on strategy, planning, research and evaluation (Toth, Serini, Wright & Emig, 2008). Concurrently, professional associations developed evaluation guidelines and toolkits for practitioners (Fairchild, 2003; Huhn, Sass & Storck, 2011; International Association for Measurement and Evaluation of Communication [AMEC], 2010; Lindenmann, 2003), with the most significant being that of the US-based Public Relations Institute (Lindenmann, 2003) which proposed three levels of evaluation: output, outtakes and outcomes. This placed communication evaluation into a similar framework (output to outcome) as the most commonly used evaluation frameworks of other fields (Mathison, 2004). With the publication by AMEC of their "Barcelona Declaration of Measurement Principles" in 2010, there has been renewed focus on evaluation and communication (Macnamara, 2014).

2.5.2. Communication evaluation models

Six main communication evaluation models have garnered attention to date: The Preparation, Implementation and Impact Model (Cutlip et al., 1994), the Pyramid Model of Public Relations Research (Macnamara, 2005), the Public Relations Effectiveness Yardstick Model (Lindenmann, 1993), the Planning, Research and Evaluation Process (Fairchild, 2003), the Unified Model of Evaluation (Watson and Noble, 2007) and the Continuing Model of Evaluation (Watson, 1997). The first three models are best viewed as taxonomies, classification schemes that categorise the effects of communications activities along a continuous scale. A limitation of taxonomies is that they do not explain the relationship between the scale items and all the elements involved, meaning their representation of reality is limited (Dubin, 1976). The last three models illustrate the steps of communication strategies and how evaluation fits within them.

The major limitation of these models is that they do not show all the relevant elements that can influence the evaluation process and therefore provide a restricted and narrow view. Consequently, models of communication evaluation developed to date do not fully meet the criteria of complete models, erring on the side of simplicity for classifications or processes and not including possible intervening elements therefore making validation and prediction difficult (Dubin, 1976; Meredith, 1992).

Building on these models and the limitations found, this study offers a more complete communication evaluation framework providing a systemic view of evaluation by incorporating the relevant elements identified beyond but linked to the evaluation process itself. This then serves as a basis for modelling and a key theoretical contribution of this study.

2.5.3. Categorising communication effects

Attempts to categorise the effects of communication activities to facilitate evaluation, as found in the above models, has been a longstanding focus of research. Lazarsfeld (1948) was one of the first to identify and categorise different levels of effects: Immediate responses, short-term effects, long-term effects and institutional change. Lazarsfeld also

made an important distinction between effects on individuals and institutions, a distinction that would emerge as key for evaluating communication campaigns some fifty years later (Coffman, 2003).

Numerous categorisations have since been proposed with most based on a hierarchical structure from producing and dissemination messages, to informing, to persuading, to inducing behaviour change on individuals or organisations (Macnamara, 2005; Rogers & Storey, 1987). These categorisations have also been adapted to fit into the standard evaluation framework of inputs, outputs and outcomes and the notions of formative, process and summative evaluation (Watson & Noble, 2007). These levels have also been matched with tools proposed to evaluate them. Levels have been added to include "inputs" that are required to produce the desired effects (Macnamara, 2005). This type of categorisation is not without its critics, coming back to its linear assumptions and appropriateness for evaluating complex programmes with unpredictable outcomes (Lennie & Tacchi, 2013; Rogers, 2008). The following table summarises where consensus has been found on the main levels of effects for communications activities with some select indicators and the tools often used for their evaluation (Cabanera-Verzosa, 2003; CCMC, 2004; Macnamara, 2005; Watson & Noble, 2007).

Table I: Categorisation of level of effects of communication activities			
Framework	**Level of effect**	**Select indicators**	**Data collection tools**
Input			
Formative evaluation	Quality and appropriateness of activity, message and channel	Level of understanding of messages by target audience	Case studies, focus groups, expert reviews, secondary data
Output			
Process evaluation	Number of communications produced, distributed and received	Percentage of target audience exposed to messages	Distribution statistics, event attendance, website visits and downloads
	Number of communications visible in the media	Number, tone and accuracy of messages in the media	Media monitoring (software or manual), web metrics software, media distribution statistics, content analysis (software or manual)

Outcome			
Summative evaluation	Changes to knowledge, attitude & behaviour	Percentage of target audience who express knowledge consistent with messages Percentage of target audience who take action as asked	Surveys, interviews, focus groups, tracking mechanisms, web metrics software
	Changes to policies, activities & practices of targeted institutions	Number and significance of change to policies of targeted institutions	Case studies, interviews, observation studies, tracking mechanisms
Impact			
Impact evaluation	Changes at national or sector level	Contribution of communications to long-term changes in society	Surveys, case studies, secondary data

Methodological developments have moved beyond these broad classifications and models, and have focused on developing methods for specific communication activities or products, including media visibility (Leinemann & Baikaltseva, 2004), employee relations (Scholes, 1997), advocacy (Patton, 2008), relationship management (Hon & Grunig, 1999) and the overall reputation of an organisation (Fonbrum, 1996). The evaluation of communication campaigns, a distinct sub-set of the broader communication activities, has received particular attention. The commonly accepted characteristics of campaigns; achieving specific outcomes within a set time period were found to be more suitable to evaluation rather than communication programmes that typically have broader objectives and run continuously or with rolling time periods (Wilson & Ogden, 2008).

A main debate on communication effects has been on the desirability to evaluate "outcomes" and "impact" over "outputs" given the significance of the former. However, in practice studies have indicated that some 80% of communication evaluations are actually at the output level (Macnamara, 2006; Watson & Noble, 2007). This study will provide an additional perspective to this debate by investigating the emphasis placed on

levels of communication effects in evaluation within IOs and any possible explanation of this choice.

2.5.4. Commonalities and prevalence of communication evaluation methodology

Given the inability to find a consensus for broader evaluation methodology, it is not surprising that there is also no consensus on a universal evaluation methodology for communication activities. Some commonality does exist in the broad methodological principles, procedures proposed and levels of effect (as seen in Table I). For example the need to define objectives of communication activities being evaluated or focusing on the effects of outcomes over outputs (AMEC, 2010; CCMC, 2004; Dozier, 1990; Fairchild, 2003; Gregory & Watson, 2008; Lindenmann, 2003; Michaelson & Macleod, 2007; Rogers & Storey, 1987).   It has been argued that no "silver bullet" solution exists (Gregory & White, 2008) with studies showing that most organisations rely on four to eight different metrics to evaluate their communication activities, pointing towards the use of multiple indicators and methods instead of a single solution (Gregory, Morgan & Kelly, 2005).   More recently, organisations have looked towards recent management trends such as Scorecards, Return on Investment calculations for solutions (Lawson, Hatch & Desroches 2007; Zerfass, 2005) and the challenge of evaluating new or social media that have become an important component of communication activities (Gregory & Watson, 2008).

Despite the broad range of methods, models, frameworks and guides produced, the actual prevalence of evaluation of communication activities is considered to be low across all sectors. Studies of prevalence estimate that between 30 - 50% of communication professionals are evaluating their programmes (private and public sector) (Fischer, 1995; Macnamara, 2006; Pohl & Vandeventer, 2001; Walker, 1994; Watson, 1997; Xavier, Patel, Johnston, Watson & Simmons, 2005).

Various elements are put forward to explain this lack of implementation and the inability of communication practitioners to implement methodology developed by academia and their own professional associations: the impracticality and complexity of methodology

required; the vagueness of communication programme design and planning making evaluation difficult; the lack of resources and know-how for evaluation; and the absence of an evaluation culture amongst communications professionals (Cutlip et al., 1994; Fischer, 1995; Kelly, 2001; Macnamara, 2006; White, 2005).

This study will provide another perspective for the methodological debate on the need or not for a universal approach for evaluating communication activities and in this regard, the appropriateness of methods used and guidance provided, in addition to assessing the prevalence of communication evaluation in an under-studied area, that is, the international public sector.

2.5.5. Communication evaluation design

By examining actual evaluations carried out of communication activities, a number of dominant trends can be seen in practical implementation, such as which evaluation designs are the most predominant. Based on an analysis of communication campaigns in the health area, Hornik (2002) found six broad categories for evaluation design: post-only, pre-post, pre-post with cohort studies, pre-post with control groups, time-series (pre, during and post) and meta-analyses. These categories are largely similar to other attempts at categorisation of evaluation designs in the communication evaluation and broader evaluation field (Broom & Dozier, 1990; Lehmann, 2007).

Hornik and others went on to study which of these designs were the most predominant in communication programmes and campaigns, finding that the either the pre-post design (without control group) or post only design were mostly used (Rogers & Storey, 1987; Hornik, 2002). For example, in an analysis of 33 health communication campaigns (Lehmann, 2007) and 34 HIV/AIDS communication campaigns (Noar, Palmgreen, Chabot, Dobransky & Zimmermann, 2009), 70% were either pre-post or post-only designs in both studies.

It has been argued that an experimental and quasi-experimental design (pre-post with control groups) is the most appropriate design when an evaluation is attempting to

determine causation and isolate the effects of the communication activities (CCMC, 2004; Dozier, 1990).

However, the use of these designs has proven challenging for communication evaluation for a number of reasons. The difficultly in assigning units (e.g. individuals or communities) randomly to control and treatment groups has been cited as a key challenge (Broom & Dozier, 1990; Flay & Cook, 1981). Another issue cited is that of "contamination" given that communication activities often seek a multiplying effect, i.e. to have publics transmit messages to other publics. The use of mass media as a communication tool often makes it impossible to "isolate" publics from potential exposure to messages (Flay & Cook, 1981).

Some authors also go further and consider the use of these designs in communication evaluation as unethical, with a scenario given of excluding a section of the public from receiving information on an urgent health issue for the sake of creating a control group (Kennedy & Abbatangelo, 2005).   In addition, questions have been raised about the capacity of communication professionals to manage designs beyond those of the post-only (Cutler 2004; Dozier, 1990).

This study will bring an additional perspective by examining the predominance of evaluation designs for communication evaluation within IOs and testing the appropriateness of these designs for this function and sector.

2.5.6. Limitation of communication evaluation literature
The above review of the literature illustrates how the communication evaluation methodology has evolved and the research that has been carried out by academia and the industry. However, several critical points can be raised on its limitations.

Discussions to date have been found to side-step the broader paradigm debates in evaluation, possibly due to the lack of epistemological and methodological foundations of communication evaluation (van Ruler, Tkalac-Verčič & Verčič, 2008). Both Dozier

(1990) and Cutler (2004) theorised that the understanding and application of appropriate methodology for communication evaluations is a major issue. Grunig (2008) argued that "metrics abound" (p. 89) but that the greatest issue is the lack of conceptualisation – the process of thinking logically and systematically about concepts, definitions and measures. As stated above, the existing models of communication evaluation lack all elements of complete conceptual modelling, limiting their potential use.

The methodology developed to date has been criticised for its lack of diversity and appropriateness (Jelen, 2008). Despite the range of guides and tools produced, the same or similar methods are proposed, such as interviews, content analysis and surveys while other more recent and innovative methods are not, such as ethnographic studies, action research, social network analysis and case studies (Grunig, 2008; Jelen, 2008).

As for the broader evaluation field, proposed evaluation approaches for communication are almost exclusively linear and logic based. In communication activities, there is a natural linearity of the actions – from identifying and researching targets, selecting activities and their implementation - which lends itself to the application of a similar linear evaluation approach. However, implementation is rarely linear as the communication activities need to adjust to how they are received and the changing environment (Manheim, 2011; van Ruler, 2015). As a result, when change happens, it is rarely predictable and often disproportionate, implying evaluation based exclusively on a linear approach is not always appropriate (Lennie & Tacchi, 2013).

The predominant theoretical models of organisational communication, such as Grunig and Hunt's (1984) Four Models of Public Relations do make reference to systems theory in their consideration of the communication process and the various interrelated influences and organisational components. However, these elements are not well projected into the communication evaluation field in the theories, methodologies and methods proposed.

The focus of communication evaluation methodology has been at the activity and

programme level, with little or no focus on developing evaluation methodologies for the other levels of potential contribution of organisational communication, i.e. at the functional (or department), organisational (or enterprise) and societal levels (Gregory, 2001; Likely & Watson, 2013). Further, in addition to these other levels, the role of relationships and intangible assets as outcomes of communication, where considerable research has been undertaken (Grunig, 2006), has been given little importance in any of the methodological guidance produced to date, such as the 2010 AMEC Principles (Likely & Watson, 2013).

Studies of communication evaluation in practice have been limited. To date, any substantial sectorial or case studies on communication evaluation have primarily been on the private sector rather than the public sector, and in North America, United Kingdom and Australia rather than other geographic regions or globally and at the programme level rather than at the organisational or societal levels (Gaunt & Wright, 2004; Gregory & White, 2008). Evaluations of campaigns has been more extensive but also limited in geographic coverage and mainly in the public health area (Coffman, 2002; Salmon, 1989).

The focus of these studies has been on the prevalence of communication evaluation, methods used and the obstacles faced, with limited connections made to the broader evaluation literature (Broom and Dozier, 1990 being one exception). As seen with the differences in the definitions of evaluation and communication evaluation (page 21), there has been little connections between the evaluation literature, theories and models and those of the communication evaluation field. As a consequence, the communication evaluation field has grown in isolation, neither addressing issues raised in the evaluation field, such as the broader paradigm debates, nor drawing from it either.

Little research has been done on these broader aspects, such as the interrelations between the evaluation process, the organisations and contexts, the management of evaluation and the use of evaluation findings. Returning to the population of interest to this study, there have been no identifiable studies on communication evaluation amongst IOs or in general at an international level (Lehmann, 1999; Sireau, 2009).

The limitations identified illustrate significant gaps in the theoretical reflections, the methodology and methods proposed and the empirical studies carried out. In this regard, the contribution of this study will be the addition of an empirical study on an understudied sector; and the development of theoretical aspects within a new conceptual framework that considers the interrelations between the evaluation process (from methodology development to use of findings), the organisations and their contexts.

2.6. References

Alkin, M. C. (Ed.). (2012). *Evaluation Roots: A Wider Perspective of Theorists' Views and Influences*. London: Sage Publications.

AMEC. (2010). *Barcelona Declaration of Measurement Principles.* Paper presented at: the 2nd European Summit on Measurement, Barcelona, 17-20 June 2010 [PDF document]. Retrieved from:
http://www.instituteforpr.org/wp-content/uploads/BarcelonaPrinciplesOct2010.pdf

Bamberger, M., Rugh, J., & Mabry, L. (2006). *Real world evaluation: Working under budget, time, data and political constraints*. Thousand Oaks, CA: Sage Publications.

Banks, S. P. (1995). *Multicultural public relations: A social-interpretive approach*. Thousand Oaks, CA: Sage Publications.

Barnett, M. N., & Finnemore, M. (1999). The politics, power, and pathologies of international organizations. *International organization*, *53*(04), 699-732.

Broom, G., & Dozier, D. (1990). *Using research in public relations*. Upper Saddle River, NJ: Prentice Hall.

Busby, J. W. (2007). Bono made Jesse Helms cry: Jubilee 2000, debt relief, and moral action in international politics. *International Studies Quarterly*, *51*(2), 247-275.

Cabanera-Verzosa, C. (2003). *Strategic Communication for Development Projects*. World Bank. Retrieved from: http://siteresources.worldbank.org/EXTDEVCOMMENG/Resources/toolkitwebjan2004.pdf

Charnovitz, S. (2006). Nongovernmental Organisations and International Law. *The American Journal of International Law*, *100*(2), 348-372.

Christie, C., & Fierro, L. (2012). Evaluation policy to implementation: An evaluation of scientifically based research in practice. *Studies in Educational Evaluation*, 38: 65–72.

Clarke, A. (1999). *Evaluation research: An introduction to principles, methods and practice*. London: Sage Publications.

Coffman, J. (2002). *Public communication campaign evaluation: An environmental scan of challenges, criticisms, practice, and opportunities*. Cambridge, MA: Harvard Family Research Project.

Coffman, J. (2003). *Lessons in evaluating communications campaigns: Five case studies*. Harvard Family Research Project. Retrieved from: http://www.hfrp.org/content/download/1114/48619/file/lessons_comm_campaigns.pdf

Communications Consortium Media Center. (2004). *Guidelines for Evaluating Non-profit Communications Efforts*. Retrieved from: http://www.researchtoaction.org/2013/03/guidelines-for-evaluating-nonprofit-communications-efforts/

Contandriopoulos, D., & Brousselle, A. (2012). Evaluation models and evaluation use. *Evaluation*, 18(1), 61-77.

Cousins, J., & Chouinard, J. (2012). *Participatory evaluation up close: A review and integration of research-based knowledge*. Charlotte, NC: Information Age Press.

Cousins, J., & Leithwood, K. (1986). Current Empirical Research on Evaluation Utilization. *Review of Educational Research*, Vol. 56, No. 3: 331-364.

Culbertson, H. M. (1996). Introduction. In H. M. Culbertson & N. Chen (Eds.), *International public relations: A comparative analysis* (pp. 1–13). Mawah, NJ: Lawrence Erlbaum.

Cutler, A. (2004). Methodical failure: The use of case study method by public relations researchers. *Public Relations Review*, 30(3), 365-375.

Cutlip, S. M., Center, A. H., & Broom, G. M. (1994). *Effective public relations*. 7th ed. Upper Saddle River, NJ: Prentice Hall.

Dahler-Larsen, P. (2012). *The Evaluation Society*. Paolo Alta, CA: Stanford Press.

Datta, L. (1994). Paradigm wars: A basis for peaceful coexistence and beyond. *New Directions for Program Evaluatio*n, 61: 53-70.

Davidson, J. (2005). *Evaluation methodology basics: the nuts and bolts of sound evaluation.* London: Sage Publications.

Deibert, R. (2000). International Plug n' Play? Citizen Activism, the Internet, and Global Public Policy. *International Studies Perspectives*, Vol. 1, No. 3: 255-272.

DeMars, W. E. (2005). *NGOs and Transnational Networks: Wild Cards in World Politics*. London: Pluto.

Dijkzeul, D., & Beigbeder, Y. (Eds.). (2003). *Rethinking international organizations: pathology and promise*. New York, NY: Berghahn Books.

Dozier, D. (1990). The innovation of research in public relations practice: Review of a program of studies. *Public Relations Research Annual*, 2, 3-28.

Dubin, R. (1976). Theory building in applied areas. In M.D. Dunnette (Ed.), *Handbook of Industrial/organizational psychology* (pp. 17-39). New York, NY: Rand McNally.

Duncan, B. (2002). Health policy in the European Union: how it's made and how to influence it. *British Medical Journal*, 324(7344): 1027–1030.

Elman, C., & Elman, M. (Eds.). (2003). *Progress in International Relations Theory*. Cambridge, MA: MIT Press.

Fairchild, M. (2003). *IPR toolkit: Media Evaluation Edition*. UK Institute for Public Relations.

Fischer, R. (1995). Control Construct design in evaluating campaigns. *Public Relations Review*, 21(1): 45-58.

Fitzpatrick, J. L. (2012). An introduction to context and its role in evaluation practice. *New Directions for Evaluation*, 2012(135), 7-24.

Fitzpatrick, J. L., Christie, C., & Mark, M. M. (2009). *Evaluation in action: Interviews with expert evaluators.* Thousand Oaks, CA: Sage Publications.

Flay, B., & Cook, T. (1981). Evaluation of Mass Media Prevention Campaigns. In R. Rice & W. Paisley, *Public Communication Campaigns* (pp. 239–264). London: Sage Publications.

Fonbrum, C. (1996). *Reputation: Realizing Value from the Corporate Image.* Boston: Harvard Business School Press.

Gaunt, R., & Wright, D. (2004). *PR Measurement*. London: Benchpoint.

Greene, J. C., Mark, M. M., & Shaw, I. (2006). *Handbook of evaluation: policies, programs and practices*. London: Sage Publications.

Gregory, A. (2001). Public relations and evaluation: Does the reality match the rhetoric? *Journal of Marketing Communication*, 7:3: 171–89.

Gregory, A., Morgan, L., & Kelly, D. (2005). *Patterns of PR in Britain's 'most admired' companies and public sector organisations*. Leeds: Leeds Business School for Eloqui Public Relations.

Gregory, A., & Watson, T. (2008). Defining the gap between research and practice in public relations programme evaluation - towards a new research agenda. *Journal of Marketing Communications*, 14:5: 337-350.

Gregory, A., & White, J. (2008). Introducing the Chartered Institute of Public Relations' initiative. In B. van Ruler, A. Tkalac-Verčič & D. Verčič (Eds.), *Public relations metrics* (pp. 307–17). New York, NY: Routledge.

Grunig, J. (Ed.). (1992). *Excellence in Public Relations and Communications Management.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Grunig, J. (2006). Furnishing the edifice: Ongoing research on public relations as a strategic management function. *Journal of Public Relations Research*, 18(2), 151–176.

Grunig, J. (2008). Conceptualizing quantitative research in public relations. In B. van Ruler, A. Tkalac-Verčič & D. Verčič (Eds.), *Public relations metrics* (pp. 88-119). New York, NY: Routledge.

Grunig, J. (2009). Paradigms of global public relations in an age of digitalisation. *PRism,* 6(2). Retrieved from:

http://www.prismjournal.org/fileadmin/Praxis/Files/globalPR/GRUNIG.pdf

Grunig, J., & Hunt, T. (1984). *Managing Public Relations*. Fort Worth, TX: HBJ Publishers.

Habermas, J. (1974). *Theory and practice*. Boston, MA: Beacon Press.

Henry, G. T., & Mark, M. M. (2003). Beyond Use: Understanding Evaluation's Influence on Attitudes and Actions. *American Journal of Evaluation*. Vol. 24, no. 3: 293-314.

Højlund, S. (2014a). Evaluation use in the organizational context – changing focus to improve theory. *Evaluation*. Vol. 20(1): 26–43.

Højlund, S. (2014b). Evaluation use in evaluation systems–the case of the European Commission. *Evaluation*, 20(4), 428-446.

Hon, L., & Grunig, J. (1999). *Guidelines for measuring relationships in public relations*. Gainesville, FL: Institute for Public Relations. Retrieved from:

http://www.instituteforpr.org/topics/measuring-relationships/

Hornik, R. (2002). *Public Health Communication. Evidence for behaviour change.* Hillsdale, NJ: LEA.

Huhn, J., Sass, J., & Storck C. (2011). *Position Paper - Communication Controlling How to maximize and demonstrate the value creation through communication*. Berlin: German Public Relations Association (DPRG). Retrieved from:

http://www.communicationcontrolling.de/fileadmin/communicationcontrolling/sonst_files/Position_paper_DPRG_ICV_2011_english.pdf

Jelen, A. (2008). The nature of scholarly endeavours in Public Relations. In B. van Ruler, A. Tkalac-Verčič & D. Verčič (Eds.), *Public relations metrics* (pp. 36-59). New York, NY: Routledge.

Johnson, K., Greenseid, L. O., Toal, S. A., King, J. A., Lawrenz, F., & Volkov, B. (2009). Research on Evaluation Use - A Review of the Empirical Literature from 1986 to 2005. *American Journal of Evaluation*, 30(3), 377-410.

Keck, M., & Sikkink, K. (1998). *Activists Beyond Borders: Advocacy Networks in international politics*. Ithaca, NY: Cornell.

Kelly, K. S. (2001). Importance and use of formal research and evaluation. In R. L. Heath (Ed.), *Handbook of public relations* (pp. 279-290). Thousand Oaks, CA: Sage Publications.

Kennedy, M., & Abbatangelo, M. (2005). *Guidance for Evaluating Mass Communication Health Initiatives: Summary of an Expert Panel Discussion* (held May 3-4, 2004). Centers for Disease Control and Prevention, Office of Communication, Atlanta, USA.

Kickbusch, I. (2003).The Contribution of the World Health Organization to a New Public Health and Health Promotion. *American Journal of Public Health*, Vol. 93, No. 3: 383-388.

Lawson, R. Hatch, T., & Desroches, D. (2007). *Scorecard best practices: design, implementation, and evaluation*. Hoboken, NJ: John Wiley and Sons.

Lazarsfeld, P. (1948). Communication research and the social psychologist. In W. Dennis, *Current Trends in Social Psychology* (pp. 218-273). Pittsburgh, PA: University of Pittsburgh Press.

Leeuw, F.L., & Furubo, J-E. (2008). Evaluation systems: what are they and why study them? *Evaluation,* 14: 157–69.

Lehmann, I. (1999). *Peacekeeping and Public Information, Caught in the crossfire*. London: Frank Cass.

Lehmann, J. (2007). *Evaluating Mass Media Health Communication*. Masters thesis (unpublished), University of Fribourg.

Leinemann, R., & Baikaltseva, E. (2004). *Media relations measurement: determining the value of PR to your company's success*. London: Gower Publishing.

Lennie, J., & Tacchi, J. (2011). *Researching, monitoring and evaluating communication for development: Trends, challenges and approached*. Report on a literature review and Consultations with Expert Reference Group and UN Focal Points on C4D. United Nations Inter-Agency Group on Communications for Development, New York, NY: UNICEF. Retrieved from:
http://www.unicef.org/cbsc/files/RME-RP-Evaluating_C4D_Trends_Challenges__Approaches_Final-2011.pdf

Lennie, J., & Tacchi, J. (2013). *Evaluating Communication for Development: A Framework for Social Change.* New York, NY: Routledge.

Likely, F., & Watson, T. (2013). Measuring the edifice: public relations measurement and evaluation practices over the course of 40 years. In K. Sriramesh, A. Zerfass & J-N. Kim (Eds.), *Public Relations and Communication Management: Current Trends and Emerging Topics* (pp. 143-162). New York, NY: Routledge.

Lindenmann, W. (1993). An 'effectiveness yardstick' to measure public relations success. *Public Relations Quarterly, 38*, 7–10.

Lindenmann, W. (2003). *Guidelines and Standards for Measuring the Effectiveness of PR Programs and Activities*. Institute for Public Relations. Retrieved from: http://www.instituteforpr.org/topics/effectiveness-programs-activities/

Macnamara, J. (1992). Evaluation: The Achilles' heel of the PR profession. *International Public Relations (IPRA) Review,* 15, No. 2: 17–31.

Macnamara, J. (2005). *Jim Macnamara's Public Relations Handbook*. Sydney: Archipelago Press.

Macnamara, J. (2006). Two-tier evaluation can help corporate communicators gain management support. *Prism* (online PR Journal), Vol. 4, Issue 2. Retrieved from: http://www.prismjournal.org/fileadmin/Praxis/Files/Journal_Files/Evaluation_Issue/COMMENTARY_MACNAMARA.pdf

Macnamara, J. (2014). The 'toe bone to the head bone' logic model to connect public relations and corporate communication to organisation outcomes. *Prism* (online PR Journal), Vol. 11, Issue 1. Retrieved from: http://www.prismjournal.org/fileadmin/11_1/Macnamara_2014.pdf

Manheim, J. (2011). *Strategy in information and influence campaigns: how policy advocates, social movements, insurgent groups, corporations, governments, and others get what they want*. New York, NY: Routledge.

Mark, M. M., Cooksy, L. J., & Trochim, W. M. (2009). Evaluation policy: An introduction and overview. *New Directions for Evaluation*, 2009(123), 3-11.

Mark, M.M., & Henry, G.T. (2004). The Mechanisms and Outcomes of Evaluation Influence. *Evaluation*. Vol. 10(1): 35–57.

Martens, K. (2002). Mission Impossible? Defining Nongovernmental Organisations. Voluntas*: International Journal of Voluntary and Nonprofit Organisations*, Vol. 13, No. 3: 271-285.

Mathison, S. (1994). Rethinking the evaluator role: Partnerships between organizations and evaluators. *Evaluation and Program Planning*, 17(3), 299-304.

Mathison, S. (2004). *Encyclopedia of evaluation.* London: Sage Publications.

Maxwell, J. A. (2012). *Qualitative research design: An interactive approach* (Vol. 41). Newbury Park, CA: Sage Publications.

Mayne, J. (2007). *Best Practices in Results-based Management: A Review of Experience: A Report for the United Nations Secretariat, Volume 1: Main Report., July 2007.* Retrieved from:
http://www.rocketsciencelab.co.uk/hipm/docs/john_mayne_un_1.pdf

McLean, A. (2000). From international organisations to international organisations. In T. Salmon & M. Imber (Eds.), *Issues in International Relations* (pp. 159-188). London: Routledge.

Meredith, J. (1992). Theory Building through Conceptual Methods. *International Journal of Operations and Production Management*, 13 (5): 3-11.

Mingst, K. (2004). *Essentials of International Relations*, 3rd ed. New York, NY: W. W. Norton and Company.

Movius, L. (2010). Cultural Globalisation and Challenges to Traditional Communication Theories, *PLATFORM: Journal of Media and Communication*, 2(1): 6-18.

Noar, S., Palmgreen, P., Chabot, M., Dobransky, N., & Zimmerman, R. (2009). 10-Year Systematic Review of HIV/AIDS Mass Communication Campaigns: Have We Made Progress? *Journal of Health Communication*, Vol. 14:15–42.

Patton, M. Q. (2008). Advocacy Impact Evaluation. *Journal of MultiDisciplinary Evaluation*, Vol. 5, no. 9.

Patton, M. Q. (2008). *Utilization-focused Evaluation,* London: Sage Publications.

Patton, M. Q. (2011). *Developmental evaluation: Applying complexity concepts to enhance innovation and use.* New York, NY: Guilford Press.

Pawson, R., & Tilley, N. (1997). *Realistic Evaluation*, London: Sage Publications.

Pohl, G., & Vanderventer, D. (2001). The workplace, undergraduate education and career preparation. In R.L. Heath (Ed.), *Handbook of public relations* (pp. 357–68). Newbury Park, CA: Sage Publications.

Picciotto, R. (2012). Experimentalism and development evaluation: Will the bubble burst? *Evaluation*, 18 (2): 213-229.

Rantanen, T. (2005). *The Media and Globalization*. London: Sage Publications.

Reeves, T. C. (1997). Established and emerging evaluation paradigms for instructional design. In Dills, C. R., & Romiszowski, A. J. (Eds.), *Instructional development paradigms* (pp. 163-178), Englewood Cliffs, NJ: Educational Technology Publications.

Reinalda, B. (Ed.). (2013). *Routledge handbook of international organization*. New York, NY: Routledge.

Reynolds, M. (2015). (Breaking) The Iron Triangle of Evaluation. *IDS Bulletin*, 46(1), 71-86.

Rogers, E., & Storey, D. (1987). Communication Campaigns. In C. Berger & H. Chaffee (Eds.), *Reader in Public Opinion and Mass Communication* (3rd ed.) (pp. 419-445). New York, NY: Free Press.

Rogers, P. J. (2008). Using programme theory for complex and complicated programs. *Evaluation*, 14(1): 29-48.

Rossi, P., Lipsey, M., & Freeman, H. (2007). *Evaluation: a systematic approach.* Thousand Oaks CA: Sage Publications.

Salmon, T. (1989). (Ed.). *Information campaigns: balancing social values and social change*. Newbury Park, CA: Sage Publications.

Salwen, M. B., & Stacks, D. W. (1996). *An Integrated approach to communication theory and research*. Hillsdale, NJ: LEA.

Sanderson, I. (2000). Evaluation in complex policy systems. *Evaluation*, 6(4), 433-454.

Scholes, E. (1997). *Gower handbook of internal communication*. London: Gower Publishing.

Schwandt, T. A. (2003). Back to the rough ground: Beyond theory to practice in evaluation. *Evaluation*, 9 (1), 353–364.

Schwandt, T. A. (2009). Globalizing influences on the western evaluation imaginary. In K.E. Ryan & J.B. Cousins (Eds.), *Sage international handbook on education evaluation* (pp. 119–39). Newbury Park, CA: Sage Publications.

Schwarz, A., & Fritsch, A. (2014). Communicating on Behalf of Global Civil Society: Management and Coordination of Public Relations in International Nongovernmental Organizations. *Journal of Public Relations Research*, 26(2), 161-183.

Scott, R. W. (1995). *Institutions and Organizations*. Newbury Park, CA: Sage Publications.

Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). Newbury Park, CA: Sage Publications.

Shadish, W. R., Cook, T. D., & Leviton, L. D. (1991). *Foundations of program evaluation*. Newbury Park, CA: Sage Publications.

Shadish, W., Newman, D, Scheirer, M., & Wye, C. (Eds.). (1995). *Guiding Principles for Evaluators*. San Francisco, CA: Jossey-Bass.

Shulha, L., & Cousins, J. (1997). Evaluation Use: Theory, Research, and Practice Since 1986. *American Journal of Evaluation,* 18: 195-208.

Simmons, P. (1998). Learning to Live with NGOs. *Foreign Policy*, *112*, 82-96.

Sireau, N. (2009). *Make Poverty History: Political Communication in Action*. Hampshire: Palgrave Macmillan.

Sriramesh, K. (2009). Globalisation and public relations: The past, present, and the future. *Prism,* 6(2). Retrieved from:
http://www.prismjournal.org/fileadmin/Praxis/Files/globalPR/SRIRAMESH.pdf

Sriramesh, K., & Verčič, D. (2007). The impact of globalization on Public Relations, *Public Relations Review*, 33(4), 355-359.

Stacks, D. (2006). *Dictionary of Public Relations Measurement and Research*. Institute for Public Relations. Retrieved from: http://www.instituteforpr.org/topics/dictionary-measurement-research/

Stern, E. (2006). Contextual challenges for evaluation practices. In I. Shaw, J. Greene & M. Mark (Eds.), *SAGE Handbook of Evaluation: Policies, Programs and Practices (pp. 292-315),* London: Sage Publications.

Tkalac, A., & Pavicic, J. (2009). Nongovernmental organizations and international public relations. In K. Sriramesh & D. Verčič. (Eds.). *The global public relations handbook: Theory, research, and practice.* 2nd ed. (pp. 807-821) New York, NY: Routledge.

Toth, E., Serini, A., Wright, D., & Emig, A. (2008). Trends in public relations roles: 1990–1995. *Public Relations Review,* Vol. 24, no. 2: 145-163.

Tourmen, C. (2009). Evaluators' decision making. *American Journal of Evaluation*, 30(1), 7-30.

Trochim, W.M.K. (2009). Evaluation policy and evaluation practice. In W.M.K. Trochim, M.M. Mark & L.J. Cooksy (Eds.), *Evaluation policy and evaluation practice. New Directions for Evaluation, 123,* 13–32.

Union of International Associations. (2014). *Yearbook of International Organizations 2014-2015, Volume 1*. Boston, MA: Brill Academic Publishers.

van Ruler, B. (2015). Agile public relations planning: The Reflective Communication Scrum. *Public Relations Review*, 41(2), 187-194.

van Ruler, B, Tkalac-Verčič, A., & Verčič D. (Eds.). (2008). *Public relations metrics.* New York, NY: Routledge.

Walker, G. (1994). Communicating public relations research. *Journal of Public Relations Research,* 6(3), 141–61.

Waltz, K. (1979). *Theory of International Politics.* New York, NY: McGraw Hill.

Watson, T. (1997). Measuring the success rate: Evaluating the PR process and PR programmes. In Kitchen, P. J. (Ed.). *Public relations: Principles and practices* (pp. 283–99). Boston, IL: International Thomson Business Press.

Watson, T., & Noble, P. (2007). *Evaluating Public Relations: a best practice guide to public relations planning, research & evaluation*, 2nd edition. London: Kogan Page.

Weiss, C.H. (1998). *Evaluation: Methods for Studying Programs and Policies* (2nd ed.). Newbury Park, CA: Sage Publications.

Weiss, C.H., Murphy-Graham, E., & Birkeland, S. (2005). An alternate route to policy influence how evaluations affect DARE. *American Journal of Evaluation*, 26(1), 12-30.

Welch, C. (2001). *NGOs and human rights: promise and performance*. Pennsylvania, PA: University of Pennsylvania Press.

Wellman, B. (2002). Designing the Internet for a Networked Society. *Communications of the ACM*, 45(5), 91-96.

White, J. (2005). *Measurement and Evaluation: Moving the Debate Forward*. London: Chartered Institute of Public Relations.

Willetts, P. (2010). *Non-governmental organizations in world politics: the construction of global governance*. London: Routledge.

Williams, B., & Hummelbrunner, R. (2010). *Systems concepts in action: a practitioner's toolkit*. Stanford, CA: Stanford University Press.

Williams, B., & Imam, I. (Eds.). (2007). *Systems Concepts in Evaluation: An Expert Anthology*. Point Reyes, CA: Edge Press/American Evaluation Society.

Wilson, L., & Ogden, J. (2008). *Strategic Communications Planning for Effective Public Relations and Marketing (5th Edition)*. New York, NY: Kendall/Hunt Publishing Company.

Xavier, R., Patel, A., Johnston, K., Watson, T., & Simmons, P. (2005). Using evaluation techniques and performance claims to demonstrate public relations impact: An Australian perspective. *Public Relations Review, 31*(3), 417–24.

3. Contribution of this thesis

3.1. A conceptual framework for IO communication evaluation

The literature review describes the research landscape and challenges relevant for evaluation in general and in particular for the communication field. The review shows that the conceptual thinking for communication evaluation and any theoretical models to date has been underdeveloped and the models limited in their potential. Further, the literature illustrates that the relations between the communication evaluation process and the various elements that can influence it is rarely considered in-depth.

Therefore, this study aims to broaden and deepen thinking and research to date in this area through the creation of a new conceptual framework for communication evaluation specific for IOs which references systems theory as its main underlying theoretical frame of reference (Figure I).

Systems theory has been seen as highly relevant to evaluation, notably in supporting the understanding of interrelationships, multiple perspectives and boundaries within the evaluation process and between it and the organisations concerned (Williams & Hummelbrunner, 2010), even if criticism has been made that it remains an abstract notion yet to be fully integrated into evaluation practice (Reynolds, 2015). Systems theory supports the standpoint taken by this author in creating this framework that organisations are made up of interrelated parts, adapting and adjusting to changes in the political, economic, and social environments in which they operate and that evaluation as process cannot be viewed in isolation of these elements (Banathy, 2000).

Systems theory was selected for this study given its broad application to organisations, their functions and sub-systems. This allowed the author to draw on the existing literature and utilise systems theory as the theoretical framework for the thesis. Other relevant theoretical concepts that could have potentially been applicable include complexity

theory, notably complex adaptive systems (Amagoh, 2008) or contingency theory (Drazin & Van de Ven, 1985). Systems theory was selected over these alternatives as it was felt to be the most comprehensive, widely used and understood.

A conceptual framework "lays out the key factors, constructs, or factors, and presumes relationships among them" (Miles and Huberman, 1994, p. 440). This author sees this framework as connecting the existing theories and research with the purpose of this study, illustrating the gaps in this body of knowledge and highlighting the original contribution the study will make. This framework has been developed through the process of philosophical conceptualisation that adds concepts and propositions to the existing body of knowledge, pulling together commonalities and patterns to offer a new perspective (Dubin, 1976; Meredith, 1992).
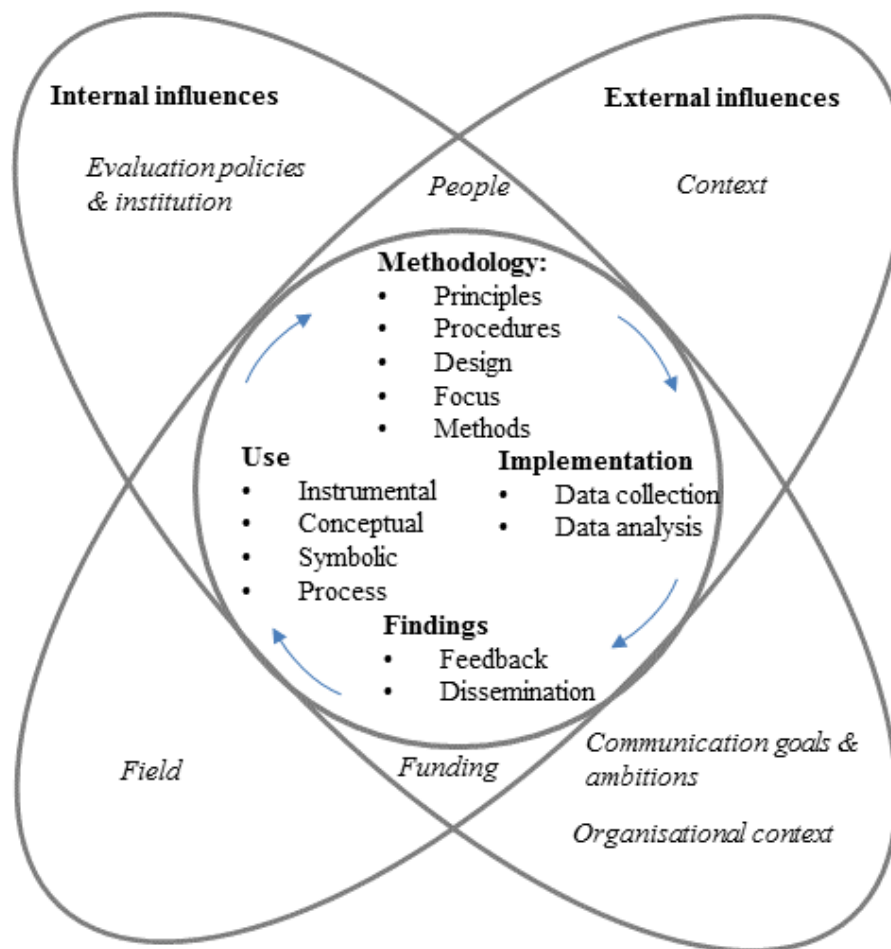
Figure III: MIFU conceptual framework for communication evaluation for IOs

At its core, the framework shows the evaluation process as four components (methodology, implementation, findings and use (MIFU)) that is cyclical rather than linear, which despite the dominance of the latter, has been seen in some evaluation frameworks, such as that of Pawson and Tilley (1997), Donaldson and Lipsey (2006) and for communication evaluation, in three of the models described previously (Fairchild, 2003; Watson and Noble, 2007; Watson, 1997). Where this framework goes beyond the existing models and frameworks is to explore further the linkages between these four components and integrates the internal and external factors that are considered to have an influence on an organisation's ability to carry out communication evaluation (as displayed in the two ellipses).

This framework purports that these four components and the internal and external factors constitutes the sphere within which communication evaluation takes place ("boundaries" in system theory). The framework highlights the interaction between the four components and the possible influence from factors outside but interlinked to the communication evaluation process. This is developed from the perspective of systems theory that recognises that organisations and their parts are interlinked and adjust to their environments through influences of both internal and external factors (that could also be labelled "constraints" or "leverage points" in systems theory (Dettmer, 2006)). "Internal" are those factors within the organisation but outside of the evaluation process; "external" are those factors outside of the evaluation process and of the organisation (Banathy, 2000; Bertalanffy, 1969).

The Methodology component comprises of the main elements considered in the choice and selection of the methodology for communication evaluation. As the literature indicates, there is no commonality on methodological elements and that the methodology used in communication evaluation is often neither appropriate nor diverse. The research will examine to what extent broad methodological principles and consequent methods and design are applied in the current practices of organisation while exploring what are the conditions necessary for communication evaluation to occur (i.e. the influence of the internal and external factors). These elements will be considered largely in the first article.

The Implementation component is concerned with the carrying out of the communication evaluation. As the literature indicates, a key challenge seen in this component is the limits to the application of methodologies, design and methods in reality, considering the internal and external factors. Research to date has not fully explored these issues and they will be considered largely in the second article.

The Findings and Use components concerns the phase of considering, communicating and using the evaluation results. Although the communication evaluation literature is limited on the issue of use, the general evaluation literature has studied extensively evaluation use, formulating the use-categorisation listed in the Use component. In the third article, this study will consider how use of communication evaluation findings occurs considering the linkages with the other two components and the internal and external factors of the framework.

The conclusions of the study will revisit the conceptual framework and draw from the findings of the previous articles on the four components and the internal and external factors, with the aim of proposing recommendations for the practice of communication evaluation within IOs.

The factors in the ellipses of the framework are described as predominantly internal or external factors, or in several cases, both (that are placed where the ellipses overlap), and are consistent with classifications and granularity of factors found in open systems frameworks (Banathy, 2000). A description of each factor is provided in the next paragraphs.

The internal factor of "communication goals and ambitions" refers to the scope of the communication activities being evaluated, such as the number and level of effects being sought (as detailed above in section 2.5.3) and the implementation models used. An implementation model could take various forms, such as hierarchical, federation, confederation, support or network-based (Brown, Ebrahim & Batliwala, 2012; Keck & Sikkink, 1998; Manheim, 2011).

The internal factor "evaluation policies and institution" refers to those policies and the institution of evaluation that is within the organisation but external to the communication functional area. The institutionalisation (or not) of evaluation and its policies across the organisation are theorised to have an impact on functional areas carrying out evaluation, such as the communication function, although it has been little studied (Bamberger, Rugh & Mabry, 2006; Christie & Fierro, 2012; Trochim, 2009).

The "organisation context" includes organisational elements such as structure, culture and strategy. Although this factor has a broad scope, the literature has identified certain organisational factors that are theorised to have an influence on communication evaluation, such as an organisation's approach to learning (strategy), its level of adaptability (culture) and the position of communication as a strategic function within the organisation (structure) (Manheim, 2011; Watson and Noble, 2007).

"People", as the professional competencies required for communication evaluation, is a factor that is categorized as both internal and external. The lack of evaluation knowledge and experience of communications staff is a well-documented issue (Cutlip, Center & Broom, 1994; Macnamara, 2006). However, it is also an external factor as the orientation of communication evaluation to date has been towards the use of external research professionals, a criticism mentioned earlier (Broom and Dozier, 1990; Watson and Noble, 2007). The human factor as an influence on the evaluation process has been previously identified and debated but mainly in reference to the various interpretations of the role played by external evaluators in the evaluation process and less so than the staff involved (Mathison, 1994; Rossi et al., 2007). Tourmen (2009) identified the demands and methodological requirements of staff on evaluation as a factor of influence.

"Funding" is also a factor categorized as both internal and external. Funding questions have consistently been found to be a barrier to communication evaluation (Cutlip et al., 1994; Macnamara, 2006; Watson and Noble, 2007). Further, funding for IOs is an external issue in that these organisations rely on outside funding for their activities,

including communication and its evaluation, an area which has been cited by one author as an issue, particularly the funding of innovative communication evaluation (Lennie & Tacchi, 2013).

The external factor of "Context" is the setting in which IOs carry out their communication activities, which the literature indicates is often complex and at the global or regional level implying an unpredictability of response to communication activities given the broad and diverse audiences being targeted (Manheim, 2011). Contextual factors are cited by some academics as key influences on the evaluation process (Pawson & Tilley, 1997) but to date have rarely been considered in-depth (Fitzpatrick, 2012).

The "Field" component is the community of organisations and bodies that have a common meaning system (peers, academia and industry associations) and the level of pressure it exerts to carry out evaluation and to which standards. Two distinct fields are relevant for this study: the general evaluation field and the communication evaluation field. It has been argued that some fields are highly demanding in terms of evaluation principles and methodologies to be applied, such as the health evaluation field (Habicht, Victora & Vaughan; 1999), whereas the research indicates that the communication evaluation field is weak in the pressure it exerts, although it has been argued that the pressure is increasing in recent years (Jelen, 2008; Likely & Watson, 2013; Watson and Noble, 2007).

In summary, the creation of a new conceptual framework for communication evaluation specific for IOs provides this study with a theoretical basis that frames this study. The framework will be operationalized in this research providing a foundation to go beyond the existing models and frameworks and explore further the linkages between the four central components of communication evaluation and the above-mentioned internal and external factors.

3.2. Contribution of this thesis by research question

As illustrated in the previous section, the conceptual framework for this study is concerned with the four components of communication evaluation, the interrelations between them and with internal and external factors. From this, the essence of the central question of this study is drawn: to what extent is communication evaluation in IOs possible? This is linked to determining what evaluation methodologies are appropriate and feasible for communication evaluation of IOs. This is further extended to considering the more systemic view, that is, the extent to which communication evaluation interacts and is impacted by internal and external factors. The findings to these points will provide a considerable contribution to the literature as described in this section with reference to the three main research questions.

> *Q1. What evaluation principles, methods and procedures are appropriate and feasible for IOs to evaluate their global communication programmes and campaigns?*

There is no known or published research into communication evaluation of IOs and this study will contribute to the limited number of empirical studies in the communication evaluation field with in-depth research of these organisations and their evaluation practices. Although the broader evaluation field has debated extensively issues of epistemological and methodological orientations and inconsistencies, this has received limited attention in the communication evaluation field, an area where this study will contribute to.

Research from the communication evaluation field has proposed principles, methods and procedures for evaluation as described above. But the understanding and application of the appropriate methodology has been highlighted as an issue and will be addressed by this study, also considering issues of designs used, levels of effects evaluated and how operating globally impacts on the methodologies used. Further, this study will provide another perspective on the need or not for a universal approach for evaluating

communication activities and in this regard, the appropriateness of methods used and guidance provided.

As part of its theoretical contribution, the study offers a more complete communication evaluation framework building and adding to the limitations found in existing models. This framework, as described above, is drawn from systems theory, providing a more holistic, systemic picture of evaluation that goes beyond what is typically researched and understood, as expanded upon in the second question:

> *Q2. To what extent do contextual, organisational and human factors influence the ability of IOs to evaluate their global communication programmes and campaigns?*

Evaluation as an isolated practice or process has been increasingly challenged in the literature as new approaches and orientations are conceptualised, disseminated and tested as described above. However, although some research has considered contextual factors and its influence on evaluation, it is rarely studied in-depth or considered in association with other possible influences, such as organisational and human factors. There has been little research into the influence of these factors on evaluation, in addition to the role of the evaluation institution, system and policies, even less so in communication evaluation and in actual studies of practices within organisations.

Through the communication evaluation framework created for this thesis, the interrelations between contextual, organisational and human factors and the evaluation process are studied and a more comprehensive perspective will be provided. Empirical research will be produced to test this model in the international public sector for the communication function which as foci are both underrepresented in research to date.

The conceptual viewpoint of evaluation as process within the larger organisation of interrelated parts and intervening factors is particularly neglected in the communication evaluation field, where this study will provide further insights. The

study will also consider the evaluation process in the broader sense, including the use of evaluation findings, as expanded upon in the third question:

Q3. How are communication evaluation findings used within IOs and what factors enable and impede their use?

An extensive body of research of both conceptual and empirical studies exists on evaluation use. However, several limitations have been identified in this body of work to which this study will make a substantial contribution. Evaluation use has typically been studied in relation to the evaluation itself, its quality, relevance and timing for example, whereas this study will investigate use in relation to broader organisational, context and human factors, where little research exists.

Further, limited research has been produced on understanding the underlying processes of how evaluation use occurs, how it is received, understood and its influence on change within organisations, with the exception of the framework produced by Henry and Mark (2003 & 2004) described above which is yet to be successfully operationalised. This study will focus on understanding these processes of use in applying the framework of Henry and Mark and producing substantive research in this area. In addition, there is no known research on evaluation use for the communication function and particularly for IOs.

3.2.1. Contribution of this thesis beyond research questions

This study will also contribute to areas beyond these specific research questions. The findings of this study can also be extended to both the non-profit and public sectors operating at the national level. For INGOs, the high majority have their origins in national-level NGOs and therefore there are many synergies between international and national level communication activities (Stroup and Murdie, 2012). For IGOs, their main national counterparts are governments and their respective ministries and there are some parallels in their communication activities as a consequence (Keohanea and Nye, 1974).

The conceptual framework of communication evaluation is designed for the international public sector. However, no complete conceptual framework or model exists that incorporates the evaluation process and the internal and external factors for communication evaluation in general. Therefore, with adaption, such a framework could be of use for broader communication evaluation beyond this specific sector.

In its totality, this study will provide theoretical and practical findings for scholarship of communication studies, social psychology, international relations and evaluation, communication professionals and those interested in the non-profit sector, nationally and internationally.   The following table provides a summary of the anticipated contribution of this study to the different fields of knowledge:

Table I: Main contribution to different fields of knowledge	
Field	**Contribution**
International relations	New findings on evaluation practices of IOs New findings on the use of evaluation findings in IOs
Organisational communications	New findings on implications for communicating globally and its evaluation
Evaluation	New findings on factors influencing the evaluation process New findings on factors influencing the use of evaluation findings New findings on intended and unintended use of evaluation findings Operationalisation of processes of use framework
Communication and social psychology	Perspectives on epistemological and methodological issues New conceptual framework for communication evaluation Critical analysis on appropriateness and feasibility of methodologies

3.3. References

Amagoh, F. (2008). Perspectives on organizational change: systems and complexity theories. *The Innovation Journal: The public sector innovation journal*, 13(3), 1-14.

Bamberger, M., Rugh, J., & Mabry, L. (2006). *Real world evaluation: Working under budget, time, data and political constraints*. Thousand Oaks, CA: Sage Publications.

Banathy, B. H. (2000). *Guided evolution of society: A systems view*. New York, NY: Kluwer Academic/Plenum Publishers.

Bertalanffy, L. (1969). *General Systems Theory: Foundations, Development, Application.* New York, NY: Braziller.

Broom, G. M., & Dozier, D. M. (1990). *Using research in public relations: Applications to program management*. Englewood Cliffs, NJ: Prentice-Hall.

Brown, L. D., Ebrahim, A., & Batliwala, S. (2012). Governing international advocacy NGOs. *World Development*, 40(6), 1098-1108.

Christie, C., & Fierro, L. (2012). Evaluation policy to implementation: An evaluation of scientifically based research in practice. *Studies in Educational Evaluation*, 38: 65–72.

Cutlip, S., Center, A., & Broom, G. (1994). *Effective public relations*. 7th ed. Upper Saddle River, NJ: Prentice Hall.

Dettmer, H. W. (2006). Systems and constraints: the concept of leverage. Port Angeles, WA: *Goal Systems International.* Retrieved January, 21, 2009.Retrieved from: http://www.goalsys.com/systemsthinking/documents/Part-6-SystemsandConstraints.pdf

Donaldson, S.I., & Lipsey, M. W. (2006). Roles for theory in contemporary evaluation practice: Developing practical knowledge. In I. Shaw, J. Greene & M. Mark, *SAGE Handbook of Evaluation: Policies, Programs and Practices (pp. 56-75),* London: Sage Publications.

Drazin, R., & Van de Ven, A. H. (1985). Alternative forms of fit in contingency theory. *Administrative science quarterly*, 514-539.

Dubin, R. (1976). Theory building in applied areas. In M.D. Dunnette (Ed.), *Handbook of Industrial/organizational psychology* (pp. 17-39). New York, NY: Rand McNally.

Fairchild, M. (2003). *IPR toolkit: Media Evaluation Edition*. UK Institute for Public Relations.

Fitzpatrick, J. L. (2012). An introduction to context and its role in evaluation practice. *New Directions for Evaluation*, 2012(135), 7-24.

Habicht, J. P., Victora, C. G., & Vaughan, J. P. (1999). Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. *International journal of epidemiology*, *28*(1), 10-18.

Henry, G. T., & Mark, M. M. (2003). Beyond Use: Understanding Evaluation's Influence on Attitudes and Actions. *American Journal of Evaluation*. Vol. 24, no. 3: 293-314.

Jelen, A. (2008). The nature of scholarly endeavours in Public Relations. In B. van Ruler, A. Tkalac-Verčič & D. Verčič (Eds.), *Public relations metrics* (pp. 36-59). New York, NY: Routledge.

Keck, M., & Sikkink, K. (1998). *Activists Beyond Borders: Advocacy Networks in international politics*. Ithaca, NY: Cornell.

Keohanea, R., & Nye, J. (1974). Transgovernmental Relations and International Organizations. *World Politics*, 27:01: 39-62.

Leinemann, R., & Baikaltseva, E. (2004). *Media relations measurement: determining the value of PR to your company's success*. London: Gower Publishing.

Lennie, J., & Tacchi, J. (2013). *Evaluating Communication for Development: A Framework for Social Change.* NY: Routledge.

Likely, F., & Watson, T. (2013). Measuring the edifice: public relations measurement and evaluation practices over the course of 40 years. In K. Sriramesh, A. Zerfass & J-N. Kim (Eds.), *Public Relations and Communication Management: Current Trends and Emerging Topics* (pp. 143-162). New York, NY: Routledge.

Macnamara, J. (2006). Two-tier evaluation can help corporate communicators gain management support. *Prism* (online PR Journal), Vol. 4, Issue 2. Retrieved from: http://www.prismjournal.org/fileadmin/Praxis/Files/Journal_Files/Evaluation_Issue/COMMENTARY_MACNAMARA.pdf

Manheim, J. (2011). *Strategy in information and influence campaigns: how policy advocates, social movements, insurgent groups, corporations, governments, and others get what they want*. New York, NY: Routledge.

Mark, M.M., & Henry, G.T. (2004). The Mechanisms and Outcomes of Evaluation Influence. *Evaluation*. Vol. 10(1): 35–57.

Mathison, S. (2004). *Encyclopedia of evaluation.* London: Sage Publications.

Meredith, J. (1992). Theory Building through Conceptual Methods. *International Journal of Operations and Production Management*, 13 (5): 3-11.

Michaelson, D., & Macleod, S. (2007). The Application of "Best Practices" In Public Relations Measurement and Evaluation Systems, *Public Relations Journal,* 1(1) 1-14.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded source book* (2nd ed.). Newbury Park, CA: Sage Publications.

Pawson, R., & Tilley, N. (1997). *Realistic Evaluation*, London: Sage Publications.

Reynolds, M. (2015). (Breaking) The Iron Triangle of Evaluation. *IDS Bulletin*, 46(1), 71-86.

Rossi, P., Lipsey, M., & Freeman, H. (2007). *Evaluation: a systematic approach.* Thousand Oaks CA: Sage Publications.

Stroup, S. S., & Murdie, A. (2012). There's no place like home: Explaining international NGO advocacy. *The Review of International Organizations*, *7*(4), 425-448.

Tourmen, C. (2009). Evaluators' decision making. *American Journal of Evaluation,* 30(1), 7-30.

Trochim, W.M.K. (2009). Evaluation policy and evaluation practice. In W.M.K. Trochim, M.M. Mark & L.J. Cooksy (Eds.), *Evaluation policy and evaluation practice*. *New Directions for Evaluation, 123,* 13–32.

Watson, T. (1997). Measuring the success rate: Evaluating the PR process and PR programmes. In P. J. Kitchen (Ed.), *Public relations: Principles and practices* (pp. 283–99). Boston, IL: International Thomson Business Press.

Watson, T., & Noble, P. (2007). *Evaluating Public Relations: a best practice guide to public relations planning, research & evaluation*, 2nd edition, London: Kogan Page.

Williams, B., & Hummelbrunner, R. (2010). *Systems concepts in action: a practitioner's toolkit*. Stanford, CA: Stanford University Press.

Interlude 1

The preceding chapters set out the research questions, a conceptual framework and the anticipated contribution of this PhD thesis. The overview of the relevant literature and research indicates the influential role of IOs and their communication yet little is known as to how these organisations manage and evaluate their activities. The evaluation literature highlights the current methodological challenges in addition to the limitations of the communication evaluation field which is in need of further theoretical contributions and empirical studies.

The thesis is described as contributing to the gap identified by considering what methodologies are appropriate and feasible for IOs to evaluate their communication activities and how is the evaluation process impacted by internal and external factors. Systems theory underlies this research with the standpoint taken that organisations such as IOs are made up of interrelated parts and evaluation, its methodologies, implementation and findings cannot be viewed in isolation.

The thesis now starts to address these issues in the next chapter, article 1. The main focus of this article is to understand how IOs have evaluated their communication activities from 1995-2010. Available evaluation reports and guidelines are reviewed systematically to assess the prevalence of evaluation within IOs and their adherence to principles of evaluation methodology.

4. Article 1 – 15 year review of communication evaluation within international organisations

Abstract

The purpose of this article is to understand how international organisations (IOs) have evaluated their communication activities and to what extent they have adhered to principles of evaluation methodology over a 15 year period (1995-2010). Forty six evaluation reports and nine guidelines from 22 organisations and four coalitions were coded on type of evaluation design and conformity with six methodology principles. Most evaluations were compliant with principle 1 (defining communication objectives), principle 3 (combining evaluation methods), principle 4 (focusing on outcomes) and principle 5 (evaluating for continued improvement). Compliance was least with principle 2 (using a rigorous design) and principle 6 (linking to organisational goals). Despite these largely positive findings, evaluation was not integrated, adopted widely or rigorously. Based on these findings, it is proposed that it is both feasible and appropriate for IOs to adopt more methodologically sound approaches through diverse methods and rigorous designs in communication evaluation. However, the ability of organisations to adopt these approaches is influenced by factors outside of the communication evaluation process, notably the integration of evaluation within the communication function and the strength of the evaluation institution and its policies, in addition to other possible organisational and contextual aspects.

Keywords: communications; communication evaluation; public relations measurement; evaluation; international organisations; non-profit communications.

4.1. Introduction

The increasing role of international organisations (IOs) in today's world has put them under the spotlight, earning them equally applause and sounding alarms (Barnett & Finnemore, 1999; McLean, 2000; Mingst, 2004; Welch, 2001). Organisations such as UNICEF, the European Union (EU), Oxfam and WWF have made increasing use of communications to profile themselves, influence issues and build relations. However, how these organisations manage and evaluate their communication activities has been little studied or analysed (Lehmann, 1999; Manheim, 2011; Sireau, 2009; Tkalac, & Pavicic, 2009).

This article sets out to partially fill this gap: To understand how IOs are evaluating their communication activities and to what extent they adhere to principles of evaluation methodology through a 15 year review of available communication evaluation reports and guidelines. This in turn provides insights as to the appropriateness (what is suitable for the organisations) and feasibility (what in practice is possible to do) of communication evaluation methodology for IOs and factors that can influence this.

Intergovernmental organisations (IGOs) are created by governments normally through international or regional treaties to undertake a variety of functions including cooperation, monitoring, dispute settlement or humanitarian intervention (Alvarez, 2006). International non-governmental organisations (INGOs) are private organisations operating in multiple countries, normally with a voluntary aspect, whose members come together for a common purpose (Mingst, 2004). INGOs undertake a variety of functions including acting as advocates for specific policies, mobilising publics, monitoring or providing humanitarian assistance. IGOs and INGOs are both considered as international non-state actors that form what is referred to as the "international public" or "international non-profit" sector (Charnovitz, 2006) and are referred to as international organisations (IOs) in this article.

The constructivist school of thought in international relations consider that IOs today are some of the most important influences on states, equal or more influential than political, military and economic actors (Barnett & Finnemore, 1999; Elman and Elman, 2003). An important difference with other actors is the new ideas and norms that these organisations produce and communicate actively on that consequently can reframe international, regional and national debates; "human rights", "climate change", "global justice" and "development goals" being prominent examples (Keck and Sikkink, 1998; Willetts, 2010). The way these organisations communicate has massively changed in the past twenty years, with organisations now more global, strategic and interactive in their communication, even if research shows their potential is yet to be reached (Grunig, 2009; Sriramesh, 2009).

To communicate, IOs use a wide variety of tools, activities and channels centralised under what is referred to as the communication function. This function within an organisation normally includes specific sub-functions such as media relations, publicity, marketing support, online communications, image/identity/reputation management and media production. Within IOs, there may be additional sub-functions related to their specific nature such as public awareness/information and education functions (Lehmann, 1999). The main action of this function and its sub-functions are communication activities, which are programmes, projects, actions and campaigns that are part of the management of communications between an organisation and its publics (Grunig, 1992).

4.2. The evolution of communication evaluation

The development of communication evaluation has been well documented with the first studies published in the 1920s as the first systematic communication activities emerged in the United States (Gregory & Watson, 2008; Watson & Noble, 2007).

In the following decades, academics concentrated on understanding how these activities functioned and what they achieved. Models and concepts of the different types of communication programmes were developed (Grunig & Hunt, 1984) and consequently

evaluation models and methodology were developed that were taught widely (Broom & Dozier, 1990; Cutlip, Center & Broom, 1994; Macnamara, 1992; Watson, 1997).

Concurrently, professional and industry associations developed guides and toolkits for practitioners (Fairchild, 2003; Huhn, Sass & Storck, 2011; International Association for Measurement and Evaluation of Communication [AMEC], 2010; Lindenmann, 2003). Latest developments have been inspired by management trends (e.g. scorecards, return on investment) and the challenge of evaluating new or social media (Chartered Institute for Public Relations, 2004; Gregory & Watson, 2008; Zerfass, 2005). Recent initiatives have also been undertaken on the specificities of communication evaluation in the non-profit sector (Communications Consortium Media Centre [CCMC], 2004).

Methodology developed to date has been found to be lacking in any epistemological and methodological foundations (van Ruler, Tkalac-Verčič & Verčič, 2008) and has not addressed the paradigm debates found in the broader evaluation field (Datta, 1994). Criticised for its lack of diversity and appropriateness (Jelen, 2008), it has been suggested that the understanding and application of appropriate methodology for communication evaluations is a major issue (Cutler, 2004; Dozier, 1990).

Criticism has been made of the theories and models developed for communication evaluation given their narrow emphasis on programme evaluation and limited consideration of broader interrelations with their organisational setting and context (Lennie & Tacchi, 2013; Likely & Watson, 2013). The predominant theoretical models of organisational communication do make reference to organisational settings and context drawing from systems theory (Broom & Dozier, 1990; Dozier, Grunig & Grunig, 2001) but this has been little applied in theoretical considerations and research on communication evaluation. As studies of communication evaluation have been criticised for their narrow isolated perspective, so too have studies on IOs for being viewed as "self contained units" (p.7) with limited attention paid to interrelations and contextual aspects (Dijkzeul & Beigbeder, 2003).

The application of existing theories and models in practice has also brought up many issues. Despite the existence of a body of guidelines and studies, the evidence indicates that the majority of companies and organisations do not systematically evaluate communication activities. Both in the private and public sector, it is estimated that between 30 - 50% of communication professionals are evaluating their programmes and with some 80% of these focusing on superficial "output" measures, such as the number of mentions in the media, (Fischer, 1995; Macnamara, 2006; Pohl & Vandeventer, 2001; Walker, 1994; Watson, 1997; Xavier, Patel, Johnston, Watson & Simmons, 2005). Reasons for this disparity include the accessibility of data to communication professionals; the impracticality and complexity of methodology required; the vagueness of communication programme design and planning making evaluation difficult; the lack of resources and know-how of evaluation; and the absence of an evaluation culture amongst communications professionals (Cutlip et al., 1994; Fischer, 1995; Kelly, 2001; Macnamara, 2006; White, 2005).

Reviews of evaluations of communication activities, similar to the review undertaken for this article, were found at the national level, notably in campaigns on influencing policy and individual behavior (Coffman, 2002; Gallagher, 1985; Weis & Tschirhart, 1994), HIV/AIDS awareness campaigns (Bertrand, O'Reilly, Denison, Anhang & Sweat, 2006; Noar, Palmgreen, Chabot, Dobransky & Zimmermann, 2009) and health campaigns (Lehmann, 2007; Synder, 2001). In general, these reviews compared the evaluation designs used and findings produced. No reviews at the international level and of IOs, the focus of this review, could be found.

4.3. Evaluation methodology for communication activities

An examination of the above-mentioned models, concepts and guides from academia, practitioners and industry associations shows no consensus on a universal evaluation methodology for communication activities. In fact, it has been argued that no "silver bullet" solution exists (Gregory & White, 2008). However, an examination of seven key references (AMEC, 2010; Broom & Dozier, 1990; CCMC, 2004; Fairchild, 2003; Lindenmann, 2003; Michaelson & Macleod, 2007; Watson & Noble, 2007) from both

academia (3) and industry (4) indicates that there is some commonality in broad methodological principles and procedures, a foundation for evaluation methodology of a given sector (Davidson, 2005). These references were selected as they represent the full scope of the current body of literature and guidance on communication evaluation methodology. They all explicitly state principles for evaluation, of which consensus is found around six points, grouped into three elements as detailed in the following table:

Table I: Presence of six principles in references						
Element:	Design		Methods	Focus		
	1. Defining objectives of communica-tion activities to be evaluated	3.Using a rigorous evaluation design	3.Using a combination of evaluation methods	4.Focusing on effects of outcomes over outputs and processes	5.Evaluating for continued improvement	6.Showing the link to overall organisational goals
Broom & Dozier	X	X	X	X	X	X
Michaelson & Macleod	X	X		X	X	X
Fairchild	X	X	X		X	X
AMEC	X	X	X	X	X	X
Lindenmann	X	X	X	X	X	X
Watson & Noble	X	X	X	X	X	X
CCMC	X	X	X	X	X	

These references also include other principles where no commonality was found amongst them, as detailed in the table below.

Table II: Additional principles featured where no consensus found						
	Applicability to broad range of activities	*Documenting methodology used*	*Procedures for specific activities, e.g. media & social media*	*Involve-ment of comm-unication staff*	*Specifying theory used*	*Considering user and situation dependent factors*
Broom & Dozier						
Michaelson & Macleod	X	X	X			
Fairchild						
AMEC			X			
Lindenmann						
Watson & Noble				X		X
CCMC				X	X	

For this review, these principles are used to understand how evaluation of communication activities of IOs has been carried out. Given their broad nature, they could be considered as minimum expectations for evaluation. However, as this review will show, on several principles, this sector has struggled to apply them. The principles are also limited in that they do not reflect all of the issues facing communication evaluation, such as the lack of theoretical foundations and diversity of methodology (Jelen, 2008; van Ruler, Tkalac-Verčič & Verčič, 2008). The limitations and gaps of the six principles are discussed at the end of this article.

To support the broader analysis of the evaluation methodology within the evaluation process and its interrelations within the organisation and context, a conceptual framework is used as seen in Figure I. Developed by this author, the framework draws from systems theory, organisational behaviour and existing empirical and theoretical research in the communication and evaluation fields (Bamberger, Rugh & Mabry, 2006; Christie & Fierro, 2012; Dubin, 1976; Fitzpatrick, 2012; Højlund, 2014a; Manheim, 2011; Meredith, 1992; Scott, 1995; Trochim, 2009; Watson & Noble, 2007).
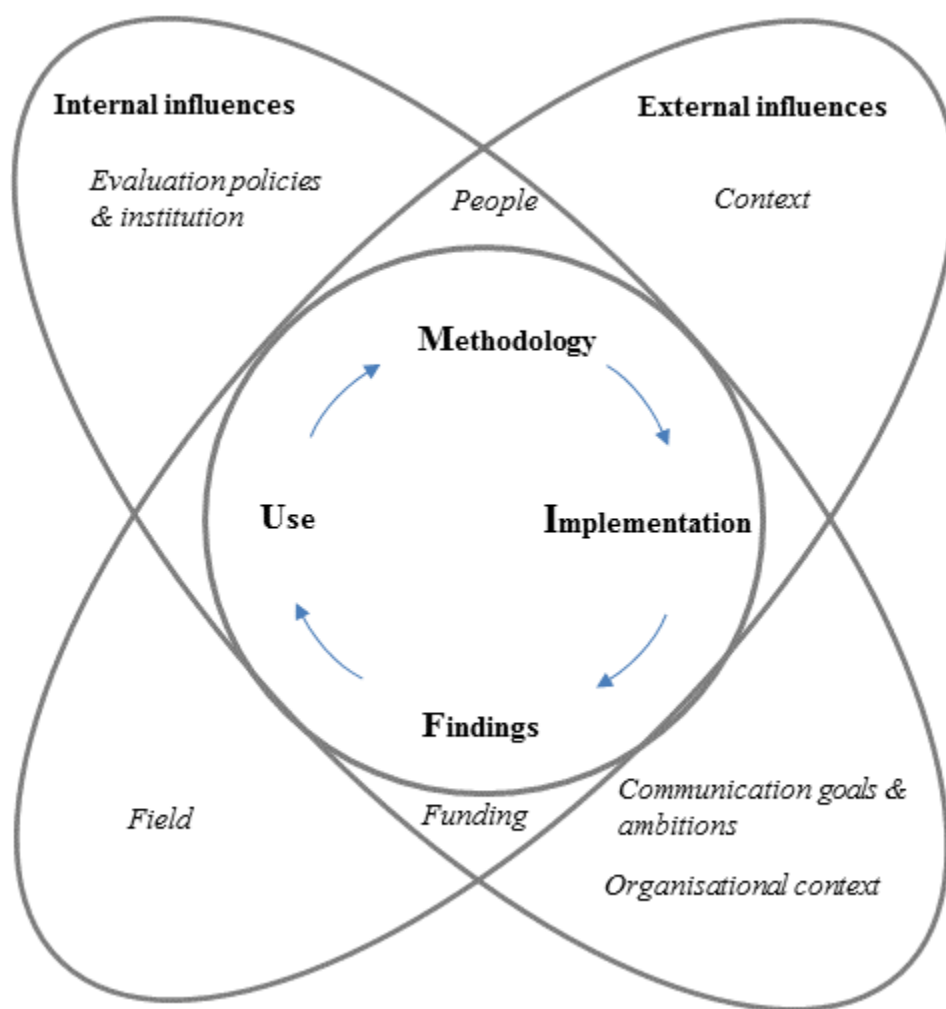
Figure IV: MIFU conceptual framework for communication evaluation for IOs

The framework highlights the interaction between the four components of the communication evaluation process (Methodology, Implementation, Findings and Use) and the possible influence from factors outside but interlinked to this process. These factors are described in Table III below based on the above-mentioned references.

Table III: Internal and external factors of the conceptual framework	
Evaluation policies and institutions	Internal to the organisation, evaluation policies, guidelines and direction as managed by a central evaluation unit; the main components of the larger evaluation system.
Organisational context	Internal contextual elements such as structure, culture and strategy.
Communication goals and ambitions	The scope of the communication activities being evaluated, such as the level of effects being sought and the

	implementation models used.
Context	The external setting within which an organisation carries out its communication activities.
Field	The community of organisations and bodies that have a common meaning system (e.g. peer organisations, academia and industry associations). Two fields are relevant for this study: evaluation and communication evaluation.
Funding	The financial resources available to communication evaluation.
People	The competencies of the human resources implicated in communication evaluation (e.g. communication staff, evaluation staff and external evaluation consultants).

4.4. Method: database and coding of IOs

4.4.1. Database and coding

This review was based on available evaluation reports and guidelines on communication activities of IOs. The review was guided by the standard protocol and stages of a systematic review: 1) development of review question and boundaries; 2) development of review protocol; 3) comprehensive search; 4) application of inclusion criteria; 5) quality assessment; 6) data extraction; and 7) synthesis of findings (Harden & Thomas, 2005; Petticrew & Roberts, 2006). Strictly speaking, the corpus studied under a systematic review would be scientific evidence (i.e. academic peer-reviewed studies) whereas the corpus of this review was made up of evaluation reports and policy documents due to the absence of the former.

A selection was made of IOs for inclusion based on existing criteria for what constitutes an IO, notably:

INGOs: Those INGOs with category 1 (general consultative status) with the UN were included. These INGOs tend to be large established INGOs with a presence in many countries and are considered of an international nature (Simmons, 1998). As of January 2011, a total of 137 organisations corresponded to this criteria.

IGOs: Those IGOs that have received a standing invitation to participate as observers in the sessions and the work of the United Nations General Assembly were included. As of January 2011, a total of 73 IOs corresponded to these criteria.

IGOs – UN: In addition, included were those IGOs that are officially recognised entities of the UN system. As of January 2011, a total of 106 organisations responded to this criteria.

Out of these 316 IOs, 86 (43 INGOs and 43 IGOs) were excluded from the review. These organisations were excluded on the basis that they did not have any communication function (such as purely coordinating or administrative bodies) or no information could be found on them. Consequently, 230 organisations were included in the review. Relevant reports, guidelines and policies were obtained from these organisations by three means:

- Making contact (through email) with the organisations.
- Searching on the websites of the organisations.
- Searching on two online databases of evaluation reports; the resources database of the Communication Initiative Network (http://www.comminit.com) and the database of evaluation reports of the Active Learning Network for Accountability and Performance in Humanitarian Action
    (http://www.alnap.org/resources/erd.aspx).

Out of these 230 organisations, evidence of communication evaluation was found in 31 IOs including nine IOs that indicated they undertook communication evaluation but that any reports or policies were confidential. For 179 IOs, publicly available information indicated that they were not undertaking communication evaluation. Of the remaining 20 IOs, no information could be found on the state of their communication evaluation. In response to contact by email, 10 IOs provided evaluation reports and/or guidelines. The online searches produced documents from another 16 IOs or coalitions of organisations.

All reports that had to be considered for the review had to meet the following inclusion criteria:

- Reports had to be published between January 1995 to December 2010;
- The reports had to be a type of review, assessment or evaluation;
- Communication had to be the central focus of the evaluation;
- The communication activities had to be global or regional in nature.

A set of 46 reports met the above criteria and were included in corpus of this review. Nine guidelines were also considered. Each report and guideline was assigned an individual label to facilitate the study, e.g. A3 for a report and G2 for a guideline. Table IV shows the total the number of reports and guidelines found for this review.

Table IV: Reports and Guidelines included in Review				
Type of organisation	*Total no. of organisations included*	*No. of organisations / coalitions where information located*	*Reports, transcripts, reviews*	*Guidelines, policies, strategies*
INGOs	94	5	6	2
IGOs	61	3	18	2
IGOs (UN)	75	14	18	5
Coalitions	--	4	4	--
Totals:	*230*	*26*	*46*	*9*

"Coalitions" has been added as a type of organisation as the review found that four reports were of communication activities conducted by coalitions of IOs rather than one single organisation (A20, A29, A34, A46).

Table VII contains a summary of the key characteristics of the 46 evaluations. Table VIII contains a summary of the nine guidelines, policies and strategies (located at the end of this article).

4.4.2. Coding

The reports were coded on the basis of the following dimensions: Type of organisation, sector, timeline for activities evaluated, coverage, theme, channels used, components used (tactics), evaluation design, data collection methods, level of effects being measured and adherence to the six evaluation principles described above. Reports were coded by this author and an independent coder. Intercoder reliability was calculated for each dimension coded (Cohen, 1960), with percentage agreement ranging from 73% to 100% with a mean of 85%. Cohen's Kappa (which corrects for chance categorisation) ranged from 0.23 to 1.0 with a mean of 0.64. These results indicate a good agreement between the coders. Discussions between the two coders resolved any differences in the coding.

4.5. Results: compliance with evaluation principles

The 46 reports represented evaluation of 46 distinct communication activities of 22 organisations and four coalitions. Most organisations were represented once or twice with the exception being the EU which was the source of 16 reports. The majority of activities were at the global level (63%) with remaining 37% at the regional level (mainly Europe and to a lesser extent Asia). The evaluations spanned 14 sectors with the dominant sectors being social (employment, culture and welfare), humanitarian aid and agriculture. The majority of activities being evaluated ran for one or two years with the majority (37/46) of evaluations undertaken from 2004 onwards. The evaluations indicated that each activity used on average three out of the four channels identified: interpersonal, media, internet and partners. A total of 31 different communication tactics were identified by the evaluations. On average, each communication activity under evaluation utilised five tactics. The most popular tactics utilised were: websites, media relations, publications and events.

The 46 evaluations and their adherence to the six principles of evaluation methodology for communication activities are now analysed and described further.

4.5.1. Principle 1: Defining objectives of communication activities to be evaluated

The notion of defining the objectives of the activities to be evaluated is stressed in the evaluation literature (Greene, Mark & Shaw, 2006; Weiss, 1998). In evaluating communication activities it is recommended to "establish clear program, activity, strategic and tactical objectives and desired outputs, outtakes and outcomes before you begin" (Lindenmann, 2003, p. 4). Dozier (1990) went further by stating that "clearly defined quantifiable objectives must be set in terms of change or maintenance of knowledge, predispositions, and behaviour of publics" (p. 5). Of note, the majority of evaluations, 78% (36/46) of this review, were undertaken by persons external to the organisations and presumably had no input in the setting of objectives of the activities they were evaluating.

In the current review, 80% (37/46) were identified as having been able to define the objectives of the communication activities being evaluated, albeit retrospectively. Those evaluations coded as complying with this principle specifically mentioned or implied strongly what the objectives of the activities were. Those coded as not complying with this principle did not mention the objectives explicitly or implicitly. Intercoder reliability for this principle was a percentage agreement of 82% and Cohen's Kappa of 0.54.

Evaluations that concentrated mainly on outputs, 28% (13/46) were typically assessing objectives at that level, i.e. activities to be held; items to be produced or distributed. Thus, determining the objectives at this level was more straightforward. In evaluating outcomes, a focus of the majority of evaluations (see principle 4), the vagueness or absence of the objectives was an issue, as illustrated by the following excerpt from an evaluation:

> One concern with this campaign is that there were no clear internally articulated realistic objectives and therefore no explicit agreement about what the campaign was actually intended to achieve. (A31, p. 11)

4.5.2. Principle 2: Using a rigorous evaluation design

Communication evaluations have been grouped into broad categories of evaluation design, namely post-only, pre-post, pre-post with cohort studies, pre-post with control groups, time-series (pre, during and post) and meta-analyses (Broom & Dozier, 1990; Hornik, 2002; Lehmann, 2007). These categories are largely similar to those used in the broader evaluation field (Rossi, Lipsey & Freeman, 2007). Previous studies of communication evaluations have found that the most common designs used are either the pre-post design (without control group) or the post-only design (Hornik, 2002; Lehmann, 2007; Rogers & Storey, 1987). This is confirmed by this review where the majority of evaluations (30/46) were post-only designs with the remaining 16 being post-only with limited time series or pre-data (13), quasi experimental design with control groups (2) and pre-post (1).

It has been argued that a rigorous evaluation design is only possible through experimental or quasi-experimental design both in evaluation in general (Rossi et al., 2007) and in communication evaluation, particularly when the evaluation is attempting to determine causation and isolate the effects of the communication activities (CCMC, 2004; Dozier, 1990). Others have argued that the choice of evaluation design is largely guided by the questions for which answers are sought (Patton, 2011; Weiss, 1998) and alternative designs have to be considered given the difficulties of undertaking experimental or quasi-experimental designs in communication evaluation. Specifically, difficulties in these designs include issues with randomly assigning units (e.g. individuals or communities) to control and treatment groups (Broom & Dozier, 1990; Flay & Cook, 1981) and "contamination" given that communication activities often seek a multiplying effect making it impossible to "isolate" publics from potential exposure to messages (Flay & Cook, 1981). Dozier (1990) further contends that the relative complexity of evaluation has led to the use of less complex applications of social science research methods to evaluate communication activities.

There is general agreement that an evaluation design needs to provide a level of rigour that gives communication managers confidence in the findings (Dozier, 1990) or

"rigorous enough to provide usable and actionable information" (Michaelson & Macleod, 2007, p. 6). Advocating for more rigorous design is also found in the available guidelines of these organisations (G8 – UNICEF; G9 – Oxfam), industry guides (Lindenmann, 2003; White, 2005) and academic research (Grunig, 2008; Jelen, 2008), often linked to promoting diversity in the methods used. For this review, "rigorous enough" was considered as the presence of comparable and varying sources of evidence to justify the conclusions made (Guba & Lincoln, 1989). On this basis, just over half of the evaluations (25/46) were considered of "a rigorous enough" design. Those evaluations coded as complying with this principle showed the use of comparable and varying sources of evidence to justify the conclusions made and an appropriate level of rigour for what they intended to evaluate (i.e. the set objectives). Those coded as not complying with this principle had limited sources of evidence and an insufficient level of rigour for what they intended to evaluate. Intercoder reliability for this principle was a percentage agreement of 73% and Cohen's Kappa of 0.44.

Those considered "rigorous enough" utilised evaluation designs appropriate for what they intended to evaluate. For example, A8 centred on evaluating quality of campaign design and implementation (output effects) and states that it is not within the scope of the evaluation to undertake an outcome or impact-focused evaluation. So in this case, the evaluation design is considered rigorous enough to evaluate the given level of effects as was the case in 11 other evaluations (A4, A7, A9, A11, A13, A14, A17, A23, A40, A41, A42).

Of those assessed as being not of a rigorous design, the majority (A3, A12, A15, A20, A24, A25, A31, A32, A38, A39, A43) lamented the limits of their evaluation designs, notably the access to, or inability to collect data needed to respond to the set evaluation questions, as illustrated by this quote from an evaluation:

> The absence of any initial appraisal of awareness/knowledge levels of segments of the general public, no matter how small these may have been, makes it

extremely difficult to pass a well-substantiated judgement on project completion, on the actual changes that the project brought about. (A24, p. 17)

In general, the evaluations steered away from questions of validity and reliability. Several (A21, A24, A30, A37, A44) mention difficulties to rule out plausible rival explanations to the changes seen, a question of internal validity (Cook & Campbell, 1979). In terms of external validity, the evaluations overall were not claiming to generalise their findings beyond the activities under review, aside from proposing recommendations for future similar activities (as described in Principle 5 below).

4.5.3. Principle 3: Using a combination of evaluation methods

A combination of quantitative and qualitative methods is advocated for widely in evaluation (Chen, 2005; Greene & Benjamin, 2001; Voils et al., 2008; Weiss, 1998; White, 2008). In communication evaluation, it is widely implied or stated directly that a combination of methods is preferred (CCMC, 2004; Lindenmann, 2003). It has also been proposed that communication practitioners fail to fully utilise the diversity of methods available or even understand how they might be applied (Jelen, 2008; Xavier et al., 2006).

In the current review, 91% (41/46) reported using more than one evaluation method with four using only one method and one evaluation (A26) not specifying the method(s) used. Although 13 methods were identified, more innovative or recent methods were largely absent, such as action research, social network analysis and ethnographic studies. Those evaluations coded as complying with this principle mentioned using more than one method. Those coded as not complying with this principle mentioned using only one method. Intercoder reliability for this principle was a percentage agreement of 100% and Cohen's Kappa of 1.0. On average, most evaluations used three methods with the most often used being interviews, document reviews and surveys. Figure II illustrates the frequency of methods used in the evaluations.
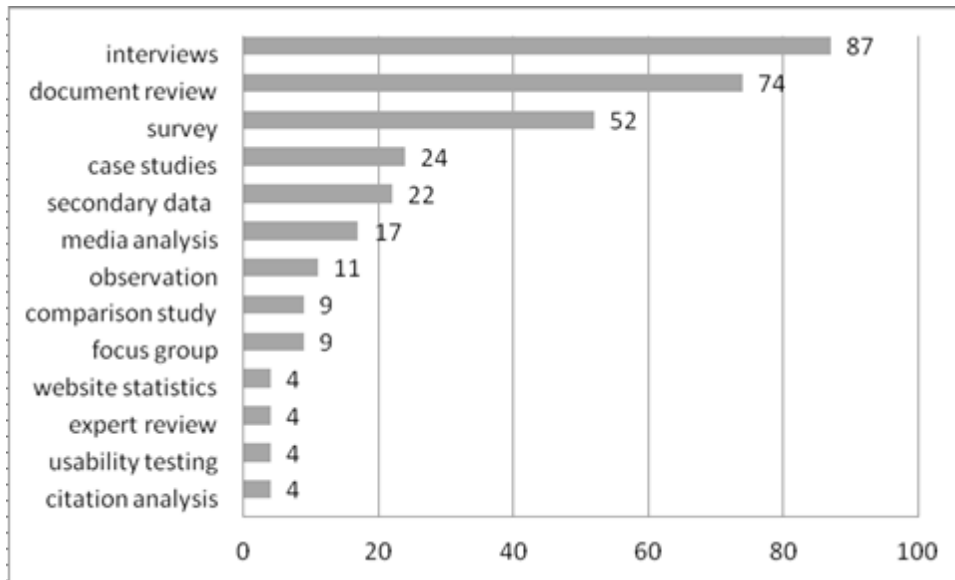
Figure II: Evaluation methods used in the 46 evaluations (%)

4.5.4. Principle 4: Focusing on effects of outcomes over outputs and processes

Evaluation of communication activities can be differentiated by the level of effect being evaluated: output (immediate effect and processes), outcome (effects on audiences and organisations) and impact (long-term effect on society or sector) (Lindenmann, 1993; Valente, 2001). These levels are similar to those used in evaluation in general (Guba and Lincoln, 1989; Weiss, 1998).

The literature advocates for a preference to the outcome level on the basis of it being of more significance and value to organisations over outputs (too superficial) and impact (too difficult to measure) (Lindenmann, 2003; Rogers and Storey, 1987). The seven principles adopted by the communication evaluation industry in 2010 dedicated one of the principles to this point, as they put simply: "Measuring the effect on outcomes is preferred to measuring outputs" (AMEC, 2010, p. 4). This focus on outcomes is also encouraged in the available guidelines on communication evaluation (G5 – World Bank; G8 – UNICEF; G9 - Oxfam). As stated above, studies of current practices indicate that up to 80% of evaluation undertaken focuses on outputs, often by examining coverage received in the media. Various reasons are put forward for this gap between theory and

practice, notably cost, ease of use, accessibility of data and complexity of undertaking outcome evaluation (Gregory & Watson, 2008; Macnamara, 1992).

Those evaluations coded as complying with this principle reported evaluating a level of effects corresponding to the outcome level. Those coded as not complying with this principle reported evaluating a level of effects corresponding to the output level. Intercoder reliability for this principle was a percentage agreement of 73% and Cohen's Kappa of 0.23.

In the current review, the majority of evaluations had some focus on outcomes (33/46). Most of these evaluations (31/46) considered more than one level of effect, with the combination of outcomes and outputs being the most frequent (28/46). The evaluation of outputs centred more on questions on efficiency and processes (19) rather than media coverage (8) differing from the trend seen in other sectors (Watson & Noble, 2007). Media sentiment (tone of coverage) was only assessed in one evaluation (A7). The evaluation of outcomes focused on effectiveness (20) (whether communication activities achieved their stated aims or not, often working through partners and relays), changes to knowledge and attitudes (14), changes to policies, activities and practices of targeted institutions (13) and changes to behaviour (10). The evaluation of impact (12) centred on estimating broader changes to sectors or society. Of significance, is the absence of evaluating relationships and other intangible assets from the evaluations, even though considerable research on their potential (and importance) as outcomes has been carried out (Grunig, 2006). There was a recognition within the evaluations of the importance of evaluating outcomes (even if the data and methods did not fully allow it as mentioned in principle 2 above), as the following extract from an evaluation illustrates:

> It is undoubtedly useful to know how many people attended a briefing or how many copies of a brochure were distributed, but it is even more useful to know what were the effects on the understanding and attitudes of those who attended the event or on those who read the document. (A17, p. 54)

4.5.5. Principle 5: Evaluating for continued improvement

Evaluation in general has moved from the notion of concentrating only on establishing whether an activity has achieved its objectives to the complementary notion of how the activity can be improved (Weiss, 1998; Wholey, 1994). In communication evaluation, Michaelson and Macleod (2007) stated, "The goal of a measurement and evaluation programme is not to determine the success or failure of a public relations programme. The goal is to improve the overall performance of these efforts" (p. 11). The Guidelines for Evaluating Non-profit Communications Efforts (2004) emphasised "Assessing whether a campaign caused its intended impact is often important … but evaluation for purposes of learning and continuous improvement is also important" (p. 3).

Those evaluations coded as complying with this principle included reflections and/or recommendations for future communication activities. Those coded as not complying with this principle did not include any reflections and/or recommendations. Intercoder reliability for this principle was a percentage agreement of 100% and Cohen's Kappa of 1.0.

In the current review, 96% (44/46) of the evaluations had a partial focus on continuous improvement. In only two evaluations (A26 and A34) there was no emphasis on continued improvements; these evaluations concentrated only on assessing the progress towards achieving objectives.

4.5.6. Principle 6: Showing the link to overall organisational goals

Communication managers often struggle to show how their activities contribute to the overall goals of their organisations (Dozier & Broom, 1995; Watson & Noble, 2007; Zerfass, 2008) and establishing such a link has been identified as a key challenge for them (Hon, 1997; Huhn et al., 2011; Macnamara, 2014). This challenge is reflected in communication evaluation where it is advocated that evaluation be designed to assess how communications' achievements contribute to the organisation as a whole (Lindenmann, 2003; Michaelson & Macleod, 2007).

Those evaluations coded as complying with this principle considered organisational goals in the findings they presented, notably by illustrating the level of support provided by communication activities to these goals and/or the links between the two. Those coded as not complying with this principle did not make any explicit link to organisational goals. Intercoder reliability for this principle was a percentage agreement of 82% and Cohen's Kappa of 0.64.

In the current review, 63% (29/46) of the evaluations endeavoured to link their findings to the overall organisational goals. For example, in A21, the evaluation had as one of its key evaluation questions how communication activities contributed to the organisational goals; in A37, the evaluation assesses the links between the communication activities and the organisation's identity. Those evaluations (17/46) that did not show the link to overall organisational objectives fell into three categories: Those that focused on evaluating communication activities; those evaluating specific communication tools, such as websites or publications; and those evaluating activities that were established to achieve objectives distinct from that of the organisation, for example, communication activities of a coalition around the 2004 Tsunami response (A34).

4.5.7. Overall compliance of the methodology principles

Overall, most evaluations were compliant with principle 3 (combination of evaluation methods) and principle 5 (evaluating for continued improvement). Compliance was least with principle 2 (use of a rigorous design) and principle 6 (link to organisational goals), as illustrated in Table V.

Table V: Overall compliance of 46 evaluations to the methodology principles					
Design		*Methods*	*Focus*		
1.Defining communication objectives	2.Rigorous design	3.Combination of evaluation methods	4.Focus on outcomes	5.Continued improvement	6.Link to organisational goals
80%	54%	91%	71%	96%	63%

Compliance to the evaluation principles can be visualised in the form of a Guttman scale, a process to determine to what extent a series of items conform to a specified set of criteria (McIver & Carmines, 1981), as seen in Table VI, located at the end of this article. A weighted score was also calculated for each evaluation report based on their compliance which ranged from three to 21. Based on the Guttman scale (Table VI), Figure III details the number of evaluation reports compliant by the number of principles, illustrating that only nine evaluations adhered to all six principles (A2, A16, A18, A19, A28, A29, A30, A44, A45).
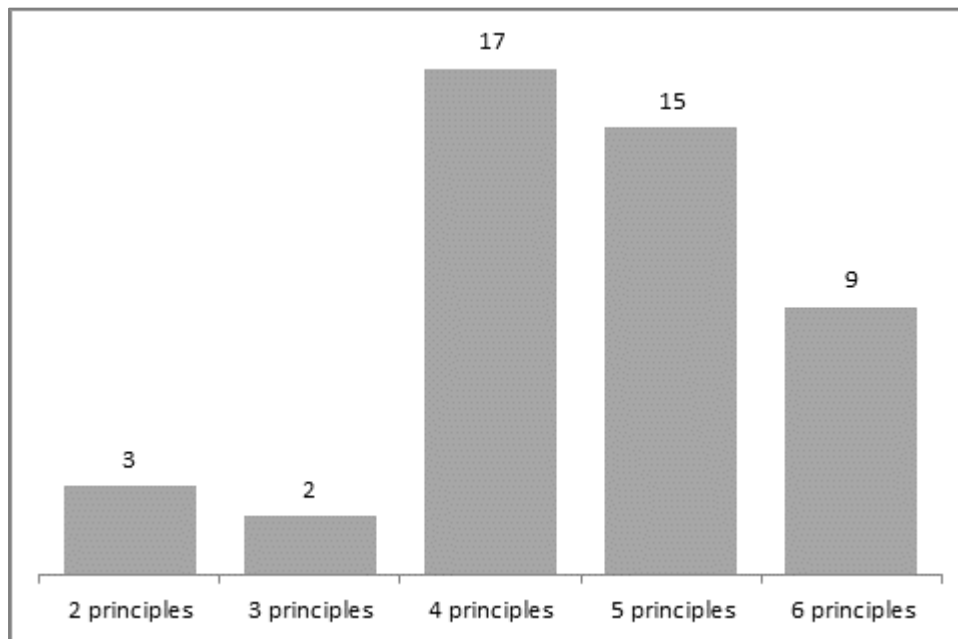
Figure III: Evaluation reports compliance with methodololgy principles (no.)

What these nine evaluations shared in common was that they were on a precise series of communication activities, e.g. campaigns/initiatives on select issues (e.g. A16 sport education, A19 food security) or precise in terms of effects desired (e.g. A2 visibility, A18 message comprehension). In addition, the evaluations were all conducted externally and the majority (7/9) were evaluation designs other than the predominant post-only design – most used some time series pre/during data and all were undertaken between 2004-2010, as were the majority of all evaluations reviewed. Overall, the type of organisation did not differ widely for the distribution and weighted score: coalitions and IGOs scored 14.75 slightly higher than INGOs at 12.33 (with median scores of 15.5, 15

and 11 respectively). The evaluations did show a gradual improvement in compliance over time, as can be seen in Figure IV mapping the average weighted score of reports by year with a linear trend line (excluding the years 1995, 1996, 1997 and 2001 where no evaluation reports were found).
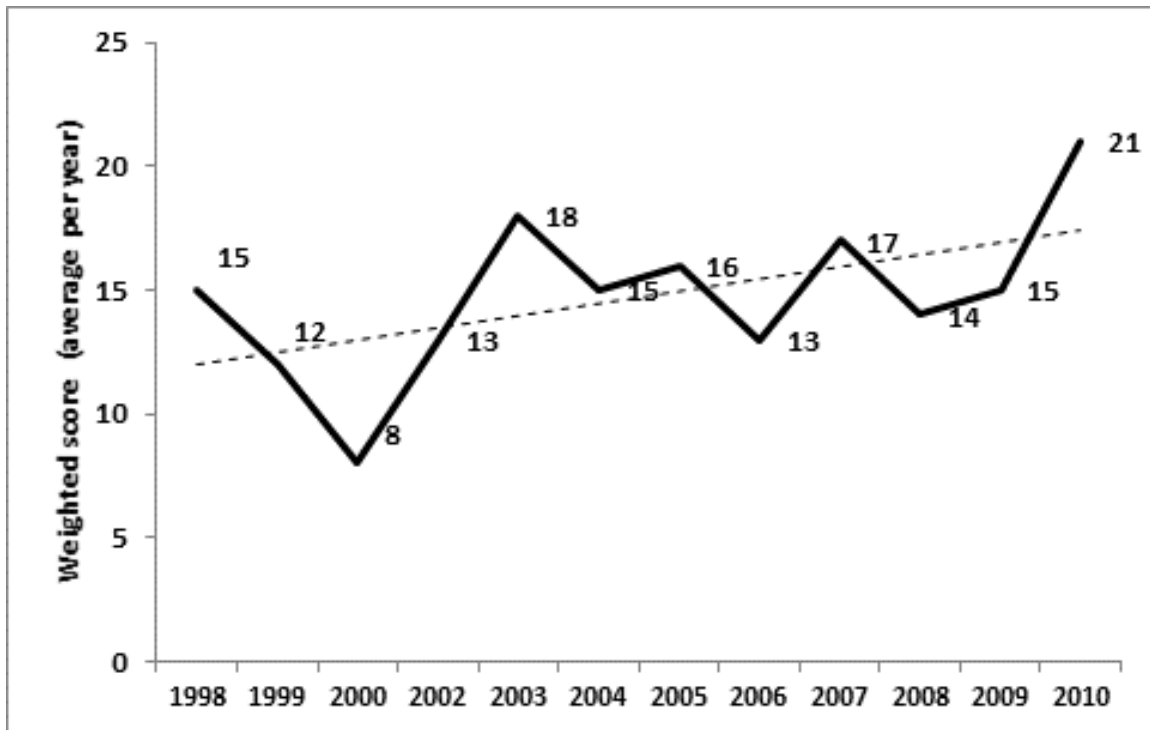
Figure IV: Average (mean) of weighted score of evaluation reports by year (no.)

As the scores of compliance decrease from six to five, the Guttman scale illustrates that adherence to principle 2 (use of a rigorous design) decreases first followed by principle 4 (focus on outcomes). From a score of five to four, adherence to principle 2 continues to decrease as does adherence to principle 6 (link to organisational goals).

Five evaluations adhered to only three (A6, A35) or two (A1, A26, A34) principles. These evaluations were unable to show a link to organisational goals (principle 6); use a rigorous evaluation design (principle 2) and three out of five used only one evaluation method (principle 3).

4.6. Discussion: understanding communication evaluation in IOs

Five main findings emerged from this systematic review that can contribute to our understanding of how IOs are evaluating their communication activities:

1) Communication evaluation was found in only 13% of IOs, lower than the estimated 30-50% for other sectors.

2) Where evaluation was carried out, the large majority of evaluations were compliant with four out of six evaluation principles: Principle 1 (defining communication objectives); principle 3 (combining evaluation methods); principle 4 (focusing on outcomes) and principle 5 (evaluating for continued improvement). Compliance was least with principle 2 (using a rigorous design) and principle 6 (linking to organisational goals). A distinction seen from other sectors was the focus on outcomes.

3) The evaluation of communication activities in IOs is largely a post-activity action that is carried out by consultants external to the communication function and the organisations.

4) The evaluations that showed greatest adherence to the evaluation principles evaluated a precise series of communication activities or level of effects and used evaluation designs other than a post-only design.

5) Compliance to the evaluation principles has increased gradually over time.

These findings are further expanded upon in the next sections.

4.6.1. Understanding low prevalence of communication evaluation

The evaluation reports directly mention obstacles to evaluation that could partially explain the low prevalence, such as access to and ability to collect data and vagueness of communication programme design, also considering that these were also found as reasons for low prevalence in other sectors (Cutlip et al., 1994; Fischer, 1995; Kelly, 2001;

Macnamara, 2006; White, 2005). It is possible that prevalence was also higher than the systematic review indicates considering the emphasis placed on monitoring in the available guidelines on communication evaluation (G2 – IAEA; G6 – FAO; G9 - Oxfam).

Where this study can bring further understanding to the issue of prevalence is by considering the IO where evaluation was predominant in this study: EU bodies. Taking into account the relative size of the EU bodies and the communication activities they produce, it is proposed that high prevalence is largely due to the strength of the evaluation institution and policies within the EU (European Commission, 2000). Put in place from the mid-1990s onwards, these policies state that evaluation is a requirement for all major EU activities and sets out the frequency with which evaluation should occur. In the available guidelines of other IOs, evaluation is encouraged (e.g. G6-FAO, G9-Oxfam) but it is not a requirement for communication activities. Organisations may also have evaluation policies in place but they may not extend to all communication activities because they are below the threshold requiring evaluation. For example, the evaluation guidelines of the International Labour Office (ILO) require an independent evaluation only for projects with a budget higher than one million US dollars (ILO, 2013).

This implies that the influence of evaluation institution and its policies is limited on communication evaluation in IOs, although it hints at its potential strong influence if it is more institutionalised and regulatory in nature given the example of the EU. This potential is supported by a study of evaluation within EU, where it was found that the institutionalisation of evaluation was largely a positive influence in making evaluation practice more systematic and routine throughout the organisation and its different functions (Højlund, 2014b). The corpus studied also indicates that influence of the evaluation institution and its policies may be more subtle and informal, such as on the focus of outcomes and methods selected, as discussed further below.

4.6.2. Explaining compliance with the evaluation principles
The high majority of evaluations were able to define the communication objectives they were to evaluate (principle 1). This illustrates that communication goals and ambitions

could be an enabling influence in that they provided guidance on the objectives under evaluation, signalling a close interrelation between these two elements. Further, the more precise the communication objectives, the higher overall compliance was with all methodology principles as seen in Table VI.

However, challenges were seen in the ability to match these objectives to an appropriate evaluation design, as seen with the low adherence to principle 2. The lack of rigorous design is not particular to the IO sector and has been seen across all sectors in studies to date (Cutler, 2004; Dozier, 1990). This deficiency has in the past been explained by budget limitations, the lack of know-how of communication staff, and in the case of experimental and quasi-experimental designs, issues of impracticality (Broom & Dozier, 1990; Flay & Cook, 1981). However, this review found an additional explanation in the corpus studied; that evaluation of communication activities are largely a post-activity action carried out by consultants external to the communication function and the organisations, indicating an absence of integration of evaluation within this function. This implies limited evaluation actions occur prior to activities commencing, evident by the dominance of post-only designs. This suggests a weak link between the evaluation process and the communication function on this point, possibly reflecting the organisational context, i.e. the emphasis placed on evaluation within functions such as communication. Nevertheless, a more rigorous evaluation design, notably a pre-post design (or if not possible a post-only design with time series or pre-data reconstructed) was feasible as was seen in evaluations with higher adherence to all methodology principles (see Table VI). This implies that in these cases evaluation was integrated earlier in the communication planning.

High compliance was found with principle 3, the combination of evaluation methods, illustrating its feasibility for communication evaluation of IOs. The preference of mixed methods is encouraged by the evaluation and communication evaluation fields, in the available evaluation guidelines (Table VIII) and based on the corpus it shows that the external consultants carrying the evaluations were competent in using multiple methods

even if their ability to use a diverse range of methods could be questioned, as discussed below.

More recent or innovative methods were not used, confirming previous findings on the lack of diversity of methods used in communication evaluation (Jelen, 2008; Xavier et al., 2006). The literature, industry guides and organisation guidelines also propose a broader range of evaluation methods that would in turn support more rigorous designs that were not used within this corpus, such as contribution analysis, process tracing, tracking studies, ethnographic studies, research action, social network analysis, reconstructed time-series data and propensity score matching (G8 – UNICEF; G9 – Oxfam), (Grunig, 2008; Jelen, 2008; Lindenmann, 2003; White, 2005; Xavier et al., 2006).

This inability to apply diverse and appropriate methods has been previously linked to funding limitations and the lack of knowledge of methods and the ability to apply them by communication staff (Cutlip et al., 1994; Dozier, 1990; Macnamara, 2006). These factors could partially explain the lack of diversity seen in the corpus studied. However, considering that the large majority of evaluations were carried out externally, it indicates that constraints are also elsewhere, such as the competencies of the external consultants to apply diverse methods. This limitation could also be due to other reasons as suggested in the literature, such as the external context that makes use of some methods impractical (due to dispersed or diverse audiences) or the organisational setting and evaluation institution and policies that do not encourage certain methods, for example, favouring quantitative methods over those of a participative and qualitative nature (Lennie & Tacchi, 2013; Williams & Imam, 2007). Therefore, although the use of more diverse methods would be appropriate, it may not be feasible depending upon the influence of the above or other factors.

IOs in their communication evaluation were found to be predominantly focused on the outcome rather than the output level as seen with the majority adherence to principle 4. The evaluation and communication evaluation fields strongly advocate for this focus that other sectors have failed to heed. This exception seen for IOs is possibly due to the

adoption of results-based management systems across this sector since the mid-nineties that has led to a move from the output to outcome levels in performance management and evaluation (Mayne, 2007), combined with an increasing interest in outcome level evaluation from the governments that largely fund IOs (Coffman, 2002; Perrin, 2006). It suggested that this orientation has consequently been reflected in the evaluation policies and the organisational contexts that then influenced the methodological approaches of communication evaluation.

However, the findings also indicate that the evaluation of outcomes faced challenges, such as the ability to match outcomes to appropriation evaluation designs and methods. Further, the outputs to outcomes model that is dominant in the literature is based on a linear concept that some scholars consider not always appropriate for evaluating communication activities, particularly those of a complex nature targeting multiple audiences across different cultures and countries, which is characteristic of many communication activities of IOs featured in these evaluations (Lennie & Tacchi, 2013; Williams & Imam, 2007).

Virtually all evaluations were found to focus on continuous improvement (principle 5) indicating strongly its feasibility for this sector and mirroring what is recommended by both the evaluation and communication evaluation fields (AMEC, 2010; CCMC, 2004; Fairchild, 2003; Lindenmann, 2003; Michaelson & Macleod, 2007; Weiss, 1998; Wholey, 1994). These sources indicate that this focus is now entrenched within evaluation and reflected in evaluation practice of both organisations and external evaluation consultants and thus reflected in the evaluation methodology used.

However, moving beyond the methodology component and to the broader evaluation process, what is considered more significant is the extent to which recommendations for improvement were acted upon by organisations (Patton, 2011), which was not measured by this review. Past studies, outside of communication evaluation and this sector, have found the direct implementation of evaluation recommendations to be limited (Shulha & Cousins, 1997; Henry and Mark, 2003; Mark and Henry, 2004). Therefore, it can be an

appropriate focus for evaluation methodology but its adherence would need to be measured at a later stage in the evaluation process, as seen in Figure I. This issue of evaluation use within IOs is the subject of another article by this author (article 3 of this PhD thesis).

The review found that not all evaluations were strong in showing how communication contributed to the organisations as a whole (principle 6), a phenomena seen in other sectors for communication evaluation (Huhn et al., 2011). The review indicated that this focus was appropriate for most communication evaluation but not for all, such as when evaluating activities that were part of a coalition and therefore more distant from an organisation's core goals. The feasibility to apply this principle is possibly influenced by the organisational context beyond the methodology and evaluation process, that in turn influenced the set goals and ambitions of the communication function. A phenomenon which has been previously studied and documented is the isolation of the communication function within the organisation as seen in its weak links to the management structure and lack of alignment with organisational goals (Grunig, 1992; Grunig, 2006).   The communication evaluation field, as seen in the academic literature and guidance of industry associations, is also limited in the advice it offers on this point, as it is overwhelmingly focused on the programme level and not on the contribution of communication to the organisational (or enterprise) and societal levels (Likely & Watson, 2013).

4.6.3. Influence of internal and external factors
 It is suggested that the ability of organisations to adhere to these six methodology principles for communication evaluation is influenced by internal and external factors to various degrees as illustrated in the previous section. The corpus supported the notion that the communication goals and ambitions have an important role to play. For example, it is implied that level of integration of evaluation in the communication function impacts on the rigour of the evaluation design adopted. The communication function providing guidance on the objectives to evaluate was seen as an enabling influence. Hence, for evaluation institution and policies, it is suggested that they could be a strong enabling

influence for evaluation in an organisation where it was institutionalised but neutral or weak in organisations where it was not. For the other five factors listed in Table III, the corpus hints at some of their influence, for example the weak influence of the field in influencing the diversity of methods selected or the complexity of the contexts in influencing access to audiences and data for evaluation. But the corpus studied cannot fully explain how these methodological elements adapt throughout the evaluation process and interact with these factors, and their respective levels of influence. This aspect is the subject of another article by this author (article 2 of this PhD).

4.6.4. Limitations of this review

Several limitations were identified in undertaking this review, notably that the assembled reports and guidelines do not represent the full body of communication evaluation of these organisations. For the majority of organisations covered by this review (179/230), the author was directly informed that this was not occurring or publicly available documentation and policies indicated that they were not undertaking communication evaluation. However, it is possible that communication evaluation is occurring in these organisations but it is not widely known internally, accessible or reflected in available documentation and policy. The limitation of access has also been seen in similar reviews (Coffman, 2002; Lehmann, 2007). In addition, although evaluation reports are the most visible and concrete outputs of evaluations (Weiss, 1998) they do not represent all evaluation actions within organisations, such as ongoing monitoring of communication activities (Starling, 2010).

In carrying out this review, limitations were identified in the ability of the six methodological principles to assess all key methodological points. For example, a gap found was the inability of these principles to assess the importance given to theory in the evaluations reviewed. This aspect is largely absent from the references but mentioned in the CCMC guidelines (2004). This is not theory about evaluation but theory about how a communication activity operates (Fitz-Gibbon & Morris, 1996), i.e. the expected path from activities to outcomes in which change is anticipated to occur, often detailed in a "programme logic" or "theory of change" (Coffman, 2003). The theory used could

influence the level of effects to be evaluated and the methods and design used. In a methodology, it sits between determining objectives and the selection of the evaluation design (Valente, 2001).

An examination of the ensemble of these principles indicates that the methodology and the corpus used for this review could provide an overview of how these organisations have evaluated their communication activities. However, to fully understand the evaluation process and the various intervening factors, further direct interaction between the researcher and the organisation would be needed. This is the approach taken by this author for articles 2 and 3 of this PhD thesis.

4.7. Conclusions: challenges to strengthening communication evaluation

This review set out to understand how IOs are evaluating their communication activities and to what extent they adhere to principles of evaluation methodology. It was found that communication evaluation was occurring in a minority of IOs. When evaluation was carried out, it used a narrow selection of methods and less than rigorous designs, although it focused more on outcomes compared to other sectors. Evaluation was found largely to be a post-activity action carried out by external consultants.

It is proposed that it is both feasible and appropriate for IOs to be more methodologically sound by adopting more diverse methods and rigorous designs in communication evaluation. This was already evident in a minority of evaluations of the corpus studied (9 out of 46). However, it is suggested that the ability of organisations to adopt this approach is influenced by factors such as the integration of evaluation within the communication function and the strength of the evaluation institution and its policies. It is implied that other organisational and contextual factors are also influential but further research directly with organisations would be required to confirm this assumption which is taken up in articles 2 and 3 of this PhD thesis.

Not surprisingly, the review showed that evaluations on more precise activities and specific levels of effect had the highest adherence to evaluation principles, implying that evaluability is assured when ambitions are limited. But evaluation methodology should be able to address more ambitious and complex outcomes. Their absence from the review illustrates the limitation in what aspect of communication is being put forward for evaluation. This is at odds with research that has found that organisations are increasingly strategic in their use of communication and expectations for what it can achieve (Grunig, 2009; Schwarz & Fritsch, 2014; Sriramesh, 2009), indicating that there is a gap between what is being done and what is being evaluated.

However, there are positive signs that IOs are aware of these challenges and are strengthening further their approaches to communication evaluation. Oxfam has been reflecting on how better to integrate evaluation in communication programme design (Starling, 2010) and has recently carried out a pilot evaluation across multiple communication campaigns using the process tracing method (Hutchings & Bowman, 2013). A number of organisations have carried out in-depth studies on how to evaluate their main communication activities including ActionAid (G1 – ActionAid), IOM (G3 – IOM) and Doctors without Borders (*Médecins Sans Frontières*) (M. Broughton, personal communication, September 14, 2011). The evaluations reviewed also indicated that both communication evaluation and compliance with the evaluation principles has increased over time.

A number of findings of this review have implications beyond the international public sector. A long-standing issue for communication evaluation has been its low adoption in practice, with this usually explained by cost, methodological complexity and know-how. However, this review found another potentially equal or more important cause, namely the place of evaluation in programme design, where communication evaluation has been almost exclusively undertaken as a post-activity action by external consultants and has limited integration within the communication function. Concerns in the literature about the lack of rigour, diversity and appropriateness of methodology and methods for communication evaluation were confirmed for this sector. However, new methods and

approaches are emerging and although not yet widespread, would be of interest to other sectors. The global nature of the communication activities of IOs was found to imply a range of issues for evaluation that are of interest to both communication and evaluation at the global level. The review also highlighted the interrelations between the methodological component of the evaluation process and internal and external factors such as strength of the evaluation institution and its policies and communication goals and ambitions.

Finally, given the influential role of IOs and their increasing use of communication, the review's findings of these organisations are of interest to studies of this global sector. Evaluation carried out to date does not yet reflect an appropriate level for their communication ambitions. However, if appropriate designs and diverse methods are adopted together with the factors of influence responded to, communication evaluation that is methodologically sound and of value to organisations should become more widespread.

Table VI: Guttman scale - compliance of 46 evaluations to the methodology principles

Codes	1		2		3		4		5		6		Score	Weighted Score*
	1	0	1	0	1	0	1	0	1	0	1	0		
A2	1		1		1		1		1		1		6	21
A16	1		1		1		1		1		1		6	21
A18	1		1		1		1		1		1		6	21
A19	1		1		1		1		1		1		6	21
A28	1		1		1		1		1		1		6	21
A29	1		1		1		1		1		1		6	21
A30	1		1		1		1		1		1		6	21
A44	1		1		1		1		1		1		6	21
A45	1		1		1		1		1		1		6	21
A27	1		1			0	1		1		1		5	19
A10		0	1		1		1		1		1		5	18
A37		0	1		1		1		1		1		5	18
A9	1		1		1			0	1		1		5	17
A17	1		1		1			0	1		1		5	17
A23	1		1		1			0	1		1		5	17
A42	1		1		1			0	1		1		5	17
A46	1		1		1		1		1			0	5	16
A12	1			0	1		1		1		1		5	15
A15	1			0	1		1		1		1		5	15
A20	1			0	1		1		1		1		5	15
A24	1			0	1		1		1		1		5	15
A38	1			0	1		1		1		1		5	15
A39	1			0	1		1		1		1		5	15
A43	1			0	1		1		1		1		5	15
A14		0	1		1			0	1		1		4	14
A4		0	1		1		1		1			0	4	13
A36	1			0		0	1		1		1		4	13
A3		0		0	1		1		1		1		4	12
A7	1		1		1			0	1			0	4	12
A8	1		1		1			0	1			0	4	12
A11	1		1		1			0	1			0	4	12
A13	1		1		1			0	1			0	4	12
A21		0		0	1		1		1		1		4	12
A25		0		0	1		1		1		1		4	12
A31		0		0	1		1		1		1		4	12
A40	1		1		1			0	1			0	4	12
A41	1		1		1			0	1			0	4	12
A5	1			0	1		1		1			0	4	10
A22	1			0	1		1		1			0	4	10
A32	1			0	1		1		1			0	4	10
A33	1			0	1		1		1			0	4	10
A35	1			0		0	1		1			0	3	8
A26	1			0		0	1			0		0	2	7
A34	1			0		0	1			0		0	2	7
A6	1			0	1			0	1			0	3	6
A1		0		0	1			0	1			0	2	3

Code	Organisation	Type	Year	Title	Coverage	Evaluation design	1	2	3	4	5	6
							...					

Table VII: Key characteristics of evaluation reports of 46 communication activities

Code	Organisation	Type	Year	Title	Coverage	Evaluation design	Principles of evaluation methodology[a] 1	2	3	4	5	6
A1	Care	INGO	2002	Lessons learned from CARE's Communications in the Afghanistan Crisis, Fall 2001	Global	post-only	No	No	Yes	No	Yes	No
A2	EU	IGO	2007	Evaluation of Communication, Information and Visibility Actions in Humanitarian Aid	Global	post-only	Yes	Yes	Yes	Yes	Yes	Yes
A3	EU	IGO	2006	Evaluation of the Information Policy on the Common Agricultural Policy	EU	post-only	No	No	Yes	Yes	Yes	Yes
A4	EU	IGO	2005	Study to assess communication, information and promotion programmes concerning beef and veal within the EU	EU	post-only (some pre and during data)	No	Yes	Yes	Yes	Yes	No
A5	EU	IGO	2004	Evaluation of the European Year of People with disabilities	EU, EES	post-only (some pre data)	Yes	No	Yes	Yes	Yes	No
A6	EU	IGO	2004	An Evaluation of the Information and Communication Strategy of the European Commission's DG for Employment and Social Affairs	EU	post-only	Yes	No	Yes	No	Yes	No
A7	EU	IGO	2004	Evaluation of the Information and Communication Activities of the European Commission's Directorate-General Energy and Transport	EU	post-only	Yes	Yes	Yes	No	Yes	No
A8	EU	IGO	2007	Midterm Evaluation of the Sustainable Energy Europe Campaign (2005-2008)	EU	post-only	Yes	Yes	Yes	No	Yes	No
A9	EU	IGO	2007	An Evaluation of Communication Links with SME Stakeholders	EU	post-only	Yes	Yes	Yes	No	Yes	Yes

A10	EU	IGO	2003	Evaluation of the information programme for the European citizen "the euro – one currency for Europe"	EU	post-only (some pre and during data)	No	Yes	Yes	Yes	Yes	Yes
A11	EU	IGO	2008	Evaluation of European Campaigns and Healthy workplace Initiatives 2006 & 2007	EU, EFTA	post-only	Yes	Yes	Yes	No	Yes	No
A12	EU	IGO	2002	Evaluation of the European Year of Languages 2001	EU, EEA	post-only	Yes	No	Yes	Yes	Yes	Yes
A13	EU	IGO	2005	Evaluation of Information and Communication (I&C) activities of DG TREN - ManagEnergy	Global	post-only (some during data)	Yes	Yes	Yes	No	Yes	No
A14	EU	IGO	2008	Evaluation of the European Year of Workers' mobility	EU	post-only (some pre data)	No	Yes	Yes	No	Yes	Yes
A15	EU	IGO	2006	Evaluation of ESF Information and Communication Activities	EU	post-only	Yes	No	Yes	Yes	Yes	Yes
A16	EU	IGO	2004	Evaluation of the European Year of Education through Sport	EU, EFTA	post-only (some pre data)	Yes	Yes	Yes	Yes	Yes	Yes
A17	EU	IGO	2008	Evaluation of DG Trade's communication policy, strategy and activities	Global	post-only	Yes	Yes	Yes	No	Yes	Yes
A18	FAO	IGO	2005	Evaluation of the Cross-organisational Strategy on Communicating FAO's Messages	Global	post-only (some during data)	Yes	Yes	Yes	Yes	Yes	Yes
A19	FAO	IGO	2006	Evaluation of FAO TeleFood	Global	post-only (some pre data)	Yes	Yes	Yes	Yes	Yes	Yes
A20	GCAP	Coalition	2006	The Global Call to Action Against Poverty (GCAP) - An External Review	Global	post-only	Yes	No	Yes	Yes	Yes	Yes
A21	Habitat	IGO	2005	Evaluation of the UN-Habitat Global Campaigns for Secure Tenure and Urban Governance	Global	post-only	No	No	Yes	Yes	Yes	Yes
A22	ILO	IGO	2006	Promoting Equality in Diversity:	EU	post-only	Yes	No	Yes	Yes	Yes	No

				Integration in Europe								
A23	ILO	IGO	2002	Evaluation of the InFocus Programme on Promoting the Declaration of Fundamental Principles and Rights at Work	Global	post-only (some during data)	Yes	Yes	Yes	No	Yes	Yes
A24	IOM	IGO	2002	Awareness Raising and Information Strategy on People in Need of International Protection	EU	post-only	Yes	No	Yes	Yes	Yes	Yes
A25	IOM	IGO	1999	Evaluation of IOM public information programmes	Global	post-only	No	No	Yes	Yes	Yes	Yes
A26	IPPF	INGO	2009	Advocacy indicators results, 2005 and 2008	Global	pre-post	Yes	No	No	Yes	No	No
A27	ITC	IGO	2009	ITC Client Survey 2009	Global	post-only	Yes	Yes	No	Yes	Yes	Yes
A28	IUCN	INGO	2004	The Knowledge Products and Services Study	Global	post-only	Yes	Yes	Yes	Yes	Yes	Yes
A29	MPH	Coalition	2005	Make Poverty History: Campaign evaluation	Global	post-only (some during data)	Yes	Yes	Yes	Yes	Yes	Yes
A30	Oxfam	INGO	2010	Evaluation of Oxfam GB's Climate Change Campaign	Global	post-only (some pre and during data)	Yes	Yes	Yes	Yes	Yes	Yes
A31	Oxfam	INGO	2004	Evaluation of Oxfam International's Coffee Campaign	Global	post-only	No	No	Yes	Yes	Yes	Yes
A32	Oxfam	INGO	2006	External Evaluation of Oxfam's Make Trade Fair Campaign	Global	post-only	Yes	No	Yes	Yes	Yes	No
A33	OCHA	IGO	2006	Evaluation of ReliefWeb	Global	post-only	Yes	No	Yes	Yes	Yes	No
A34	TEC	Coalition	2006	The Tsunami Evaluation Coalition Media Evaluation	Global	post-only	Yes	No	No	Yes	No	No
A35	UNCTAD	IGO	2000	Independent review: selected UNCTAD Technical Cooperation Publications	Global	post-only	Yes	No	No	Yes	Yes	No
A36	UNDG	IGO	2008	Communicating as One - Lessons Learned from Delivering as One in 2007	Global	post-only	Yes	No	No	Yes	Yes	Yes
A37	UNDP	IGO	2009	UN Millennium Campaign -	Global	post-only (some pre	No	Yes	Yes	Yes	Yes	Yes

							1	2	3	4	5	6
				external evaluation 2009		and during data)						
A38	UNEP	IGO	2002	Evaluation report - division of Communications and Public Information	Global	post-only	Yes	No	Yes	Yes	Yes	Yes
A39	UNESCO	IGO	2005	Evaluation of UNESCO's Partnerships Aimed at Strengthening Communication Capacities	Global	post-only	Yes	No	Yes	Yes	Yes	Yes
A40	UNSECSO	IGO	2005	Ensuring that Publications and Other Materials Released from UNESCO Meet Appropriate Quality Standards	Global	post-only	Yes	Yes	Yes	No	Yes	No
A41	UNESCO	IGO	2004	Evaluation of the UNESCO Communication and Information Sector's WebWorld Int. initiative	Global	post-only	Yes	Yes	Yes	No	Yes	No
A42	UNHCR	IGO	2005	"Evaluation of the Department of International Protection's Protection Information Section	Global	post-only	Yes	Yes	Yes	No	Yes	Yes
A43	UNHCR	IGO	1998	Review of UNHCR Mass Information Activities	Global	post-only	Yes	No	Yes	Yes	Yes	Yes
A44	UNICEF	IGO	2004	Evaluation of the Meena Communication Initiative	South Asia	quasi-experimental	Yes	Yes	Yes	Yes	Yes	Yes
A45	UNICEF	IGO	2005	UNICEF Tsunami Communications Evaluation	East Asia	post-only (some pre and during data)	Yes	Yes	Yes	Yes	Yes	Yes
A46	FHI	Coalition	2005	Using Global Media to Reach Youth: 2002 MTV Campaign	Global	quasi-experimental	Yes	Yes	Yes	Yes	Yes	No

Note. [a]Principles of evaluation methodology for communication activities:

1. Defining objectives of communication activities to be evaluated   2. Using a rigorous evaluation design
3. Using a combination of evaluation methods   4. Focusing on effects of outcomes over outputs and processes
5. Evaluating for continued improvement   6. Showing the link to overall organisational goals

Table VIII: Policies, guidelines and strategies – communication evaluation				
Code	*Organisation*	*Type*	*Date*	*Title*
G1	ActionAid	INGO	2001	Monitoring and Evaluating Advocacy: A Scoping Study
G2	IAEA	IGO	2010	Public Information and Communications Policy
G3	IOM	IGO	2001	Research and Evaluation Methodology for Mass Information Activities – An institutional approach
G4	UNDG	IGO	2010	UNDG/DOCO Communication Strategy 2010
G5	World Bank	IGO	2003	Strategic communication for Development Projects
G6	FAO	IGO	2004	Corporate Communication Policy and Strategy
G7	ICTP	IGO	2009	Communication Strategy 2010
G8	UNICEF	IGO	2008	Researching, Monitoring and Evaluating Strategic Communication for Behaviour and Social Change
G9	Oxfam	INGO	2010	Monitoring, Evaluating and Learning in Oxfam International Campaigns

***Explanation for the weighted score of Table VI:** A weighted score was created for the Guttmann scale based on a weighting of each principle relative to their frequency in the scale. For example, principle 2 was the least frequent in the scale, therefore it was accorded a score of six; principle 5 was the most frequent in scale; therefore it was accorded a score of 1. The column Weighted score is the total of the Weighted score per evaluation report, which are listed as individual rows in the table.

4.8. References

Alvarez, J. (2006). International Organisations: Then and Now. *The American Journal of International Law, 100*, (2), 324-347.

AMEC. (2010). *Barcelona Declaration of Measurement Principles.* Article presented at:

the 2nd European Summit on Measurement, Barcelona, 17-20 June 2010 [PDF document]. Retrieved from:

 http://www.instituteforpr.org/wp-content/uploads/BarcelonaPrinciplesOct2010.pdf

Bamberger, M., Rugh, J., & Mabry, L. (2006). *Real world evaluation: Working under budget, time, data and political constraints*. Thousand Oaks, CA: Sage Publications.

Barnett, M., & Finnemore, M. (1999). The Politics, Power, and Pathologies of International Organisations. *International Organisation*, *53*(4), 699-732.

Bertrand, J. T., O'Reilly, K., Denison, J., Anhang, R., & Sweat, M. (2006). Systematic review of the effectiveness of mass communication programs to change HIV/AIDS-related behaviors in developing countries. *Health Education Research*, *21*, 567–597.

Broom, G., & Dozier, D. (1990). *Using research in public relations*. Upper Saddle River, NJ: Prentice Hall.

Charnovitz, S. (2006). Nongovernmental Organisations and International Law. *The American Journal of International Law*, 100(2), 348-372.

Chartered Institute for Public Relations. (2004). *Best Practice in the Measurement and Reporting of Public Relations and ROI.* London: Chartered Institute for Public Relations.

Chen, H. (2005). *Practical Program Evaluation: Assessing and Improving Planning, Implementation and Effectiveness*. Thousand Oaks, CA: Sage Publications.

Christie, C., & Fierro, L. (2012). Evaluation policy to implementation: An evaluation of scientifically based research in practice. *Studies in Educational Evaluation*, 38: 65–72.

Coffman, J. (2002). *Pubic communication campaign evaluation - An environmental scan of challenges, criticisms, practice and opportunities*. Harvard Family Research Project. Retrieved from: http://www.mediaevaluationproject.org/HFRP.pdf

Coffman, J. (2003). *Lessons in evaluating communications campaigns: Five case studies*. Harvard Family Research Project. Retrieved from: http://www.mediaevaluationproject.org/HFRP2.pdf

Communications Consortium Media Center. (2004). *Guidelines for Evaluating Nonprofit Communications Efforts*. Retrieved from: http://www.mediaevaluationproject.org/Paper5.pdf

Cohen, J. A. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37–46.

Cook, T. D., & Campbell D. T. (1979). *Quasi Experimentation: Design and Analytical Issues for Field Settings.* Chicago, IL: Rand McNally.

Cutler, A. (2004). Methodical failure: The use of case study method by public relations researchers. *Public Relations Review*, 30(3), 365-375.

Cutlip, S. M., Center, A. H., & Broom, G. M. (2006). *Effective public relations* (9th ed.). Upper Saddle River, NJ: Prentice Hall.

Datta, L. (1994). Paradigm wars: A basis for peaceful coexistence and beyond. *New Directions for Program Evaluatio*n, 61: 53-70.

Davidson, J. (2005). *Evaluation methodology basics: the nuts and bolts of sound evaluation.* London: Sage Publications.

Dijkzeul, D., & Beigbeder, Y. (Eds.). (2003). *Rethinking international organizations: pathology and promise*. New York, NY: Berghahn Books.

Dozier, D. (1990). The innovation of research in public relations practice: Review of a program of studies. *Public Relations Research Annual, 2*, 3-28.

Dozier, D., & Broom, G. (1995). Evolution of the manager role in public relations practice. *Journal of Public Relations Research, 7*(1), 3-26.

Dozier, D., Grunig, L., & Grunig, J. (2001). Public Relations as Communications Campaign. In R. Rice & C. Atkins (Eds.), *Public Communication Campaigns* (3rd ed.) (pp. 231-248). Newbury Park, CA: Sage Publications.

Dubin, R. (1976). Theory building in applied areas. In M.D. Dunnette (Ed.), *Handbook of Industrial/organizational psychology* (pp. 17-39). New York, NY: Rand McNally.

European Commission. (2000). *Focus on Results: Strengthening evaluation of Commission activities, SEC (2000)1051 - 26/07/2000, Communication to the Commission from Mrs. Schreyer in agreement with Mr. Kinnock and the President, according to action 16 of the Action Plan for Reform.* Retrieved from:
http://ec.europa.eu/dgs/information_society/evaluation/data/pdf/lib_master/sec2000_1051 _strengthening_eval.pdf.

Fairchild, M. (2003). *IPR toolkit: Media Evaluation Edition*. London: Charted Institute for Public Relations.

Fischer, R. (1995). Control construct design in evaluating campaigns. *Public Relations Review*, *21*(1), 45-58.

Fitz-Gibbon, C., & Morris, L. (1996). Theory-based Evaluation. *Evaluation Practice*, Vol. 17, No. 2, 177-184.

Fitzpatrick, J. L. (2012). An introduction to context and its role in evaluation practice. *New Directions for Evaluation*, 2012(135), 7-24.

Flay, B., & Cook, T. (1981). Evaluation of Mass Media Prevention Campaigns. In R. Rice & W. Paisley, *Public Communication Campaigns* (pp. 239–264). London: Sage Publications.

Gallagher, M. (1985). *Becoming aware: human rights and the family – a study based on four communication campaigns*. Paris: UNESCO.

Greene, J. C., & Benjamin, L. (2001). The Merits of Mixing Methods in Evaluation. *Evaluation, 7*(1), 25-44.

Greene, J. C., Mark, M. M., & Shaw, I. (2006). *Handbook of evaluation: policies, programs and practices*. London: Sage Publications.

Gregory, A., & Watson, T. (2008). Defining the gap between research and practice in public relations programme evaluation - towards a new research agenda. *Journal of Marketing Communications*, *14*(5), 337-350.

Gregory, A., & White, J. (2008). Introducing the Chartered Institute of Public Relations' initiative. In B. van Ruler, A. Tkalac-Verčič & D. Verčič (Eds.), *Public relations metrics* (pp. 307–17). New York, NY: Routledge.

Grunig, J. (ed.) (1992). *Excellence in Public Relations and Communications Management.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Grunig, J. (2006). Furnishing the edifice: Ongoing research on public relations as a strategic management function. *Journal of Public Relations Research*, 18(2), 151–176.

Grunig, J. (2008). Conceptualizing quantitative research in public relations. In B. van Ruler, A. Tkalac-Verčič & D. Verčič (Eds.), *Public relations metrics* (pp. 88-119). New York, NY: Routledge.

Grunig, J. (2009). Paradigms of global public relations in an age of digitalisation. *PRism,* 6(2). Retrieved from:
http://www.prismjournal.org/fileadmin/Praxis/Files/globalPR/GRUNIG.pdf

Grunig, J., & Hunt, T. (1984). *Managing Public Relations*. Fort Worth, TX: HBJ Publishers.

Guba, E., & Lincoln, Y. (1989). *Fourth generation evaluation.* London: Sage Publications.

Harden, A., & Thomas J. (2005). Methodological Issues in Combining Diverse Study Types in Systematic Reviews. *International Journal of Social Research Methodology, 8*, 257–271.

Højlund, S. (2014a). Evaluation use in the organizational context − changing focus to improve theory. *Evaluation*. Vol. 20(1): 26–43.

Højlund, S. (2014b). Evaluation use in evaluation systems−the case of the European Commission. *Evaluation*, 20(4), 428-446.

Hon, L. (1997). What Have You Done For Me Lately? Exploring Effectiveness in Public Relations. *Journal of Public Relations Research*, 9(1), 1-30.

Hornik, R. (2002). *Public Health Communication. Evidence for behaviour change.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Huhn, J. Sass, J., & Storck, C. (2011). *Position paper - Communication Controlling How to maximize and demonstrate the value creation through communication*. Berlin: German Public Relations Association (DPRG). Retrieved from: http://www.communicationcontrolling.de/fileadmin/communicationcontrolling/sonst_files/Position_paper_DPRG_ICV_2011_english.pdf

Hutchings, C., & Bowman, K. (2013). Lessons learned from a year of advocacy impact evaluation. *Oxfam Policy and Practice Blog*. Retrieved from: http://policy-practice.oxfam.org.uk/blog/2013/02/lessons-learned-from-a-year-of-advocacy-impact-evaluation

International Labour Office. (2013). *ILO policy guidelines for results-based evaluation: principles, rationale, planning and managing for evaluations*. International Labour Office, Evaluation Unit (EVAL), second edition, Geneva.

Jelen, A. (2008). The nature of scholarly endeavours in Public Relations. In B. van Ruler, A. Tkalac-Verčič & D. Verčič (Eds.), *Public relations metrics* (pp. 36-59). New York, NY: Routledge.

Keck, M., & Sikkink, K. (1998). *Activists Beyond Borders: Advocacy Networks in international politics*. Ithaca, NY: Cornell.

Kelly, K. S. (2001). Importance and use of formal research and evaluation. In R. L. Heath (Ed.), *Handbook of public relations* (pp. 279-290). Thousand Oaks, CA: Sage Publications.

Lehmann, I. (1999). *Peacekeeping and Public Information, Caught in the crossfire*. London: Frank Cass.

Lehmann, J. (2007). *Evaluating Mass Media Health Communication*. (Unpublished master's dissertation). University of Fribourg, Fribourg, Switzerland.

Lennie, J., & Tacchi, J. (2013). *Evaluating Communication for Development: A Framework for Social Change.* NY: Routledge.

Likely, F., & Watson, T. (2013). Measuring the edifice: public relations measurement and evaluation practices over the course of 40 years. In K. Sriramesh, A. Zerfass & J-N. Kim (Eds.), *Public Relations and Communication Management: Current Trends and Emerging Topics* (pp. 143-162). New York, NY: Routledge.

Lindenmann, W. (1993). An 'effectiveness yardstick' to measure public relations success. *Public Relations Quarterly, 38*, 7–10.

Lindenmann, W. (2003). *Guidelines and Standards for Measuring the Effectiveness of PR Programs and Activities*. Gainesville, FL: Institute for Public Relations. Retrieved from: http://www.instituteforpr.org/iprwp/wp-content/uploads/2002_MeasuringPrograms.pdf

Lindenmann, W. (2006). *Public Relations Research for Planning and Evaluation.* Gainesville, FL: Institute for Public Relations. Retrieved from: http://www.instituteforpr.org/wp-content/uploads/2006_Planning_Eval2.pdf

Macnamara, J. (1992). Evaluation: The Achilles' heel of the PR profession. *International Public Relations Review, 15*(2), 17–31.

Macnamara, J. (2006). Two-tier evaluation can help corporate communicators gain management support. *PRism*, *4*(2). Retrieved from: http://www.prismjournal.org/fileadmin/Praxis/Files/Journal_Files/Evaluation_Issue/COMMENTARY_MACNAMARA.pdf

Macnamara, J. (2014). The 'toe bone to the head bone' logic model to connect public relations and corporate communication to organisation outcomes. *Prism* (online PR Journal), Vol. 11, Issue 1. Retrieved from:

http://www.prismjournal.org/fileadmin/11_1/Macnamara_2014.pdf

Manheim, J. (2011). *Strategy in information and influence campaigns: how policy advocates, social movements, insurgent groups, corporations, governments, and others get what they want*. New York, NY: Routledge.

Mark, M., & Henry, G. (2003). Beyond Use: Understanding Evaluation's Influence on Attitudes and Actions. *American Journal of Evaluation*. Vol. 24, no. 3: 293-314.

Mark, M., & Henry, G. (2004). The Mechanisms and Outcomes of Evaluation Influence. *Evaluation*. Vol. 10(1): 35–57.

Mayne, J. (2007). *Best Practices in Results-based Management: A Review of Experience: A Report for the United Nations Secretariat, Volume 1: Main Report., July 2007.* Retrieved from:
http://www.rocketsciencelab.co.uk/hipm/docs/john_mayne_un_1.pdf

Meredith, J. (1992). Theory Building through Conceptual Methods. *International Journal of Operations and Production Management*, 13 (5): 3-11.

McIver, P., & Carmines, E. G. (1981). *Unidimensional scaling* (Sage University Paper Series on Quantitative Applications in the Social Science, series no. 24). Newbury Park, CA: Sage Publications.

McLean, A. (2000). From international organisations to international organisations. In T. Salmon & M. Imber (Eds.), *Issues in International Relations* (pp. 159-188). London: Routledge.

Michaelson, D., & Macleod, S. (2007). The Application of "Best Practices" In Public Relations Measurement and Evaluation Systems, *Public Relations Journal, 1*(1) 1-14.

Mingst, K. (2004). *Essentials of International Relations* (3rd ed.). New York, NY: W. W. Norton and Company.

Noar, S., Palmgreen, P., Chabot, M., Dobransky, N., & Zimmermann, R. (2009). A 10-Year Systematic Review of HIV/AIDS Mass Communication Campaigns: Have We Made Progress? *Journal of Health Communication*, *14*, 15–42.

Patton, M. Q. (2011). *Developmental evaluation: Applying complexity concepts to enhance innovation and use.* New York, NY: Guilford Press.

Perrin, B. (2006). *Moving from Outputs to Outcomes: Practical advice from governments around the world*. Managing for performance and results series, Washington, DC: World Bank. Retrieved from:
 http://siteresources.worldbank.org/CDFINTRANET/Resources/PerrinReport.pdf

Petticrew, M., & Roberts, H. (2006). *Systematic Reviews in the Social Sciences: a Practical Guide*. Oxford: Blackwell Publishing.

Pohl, G., & Vanderventer, D. (2001). The workplace, undergraduate education and career preparation. In R.L. Heath (Ed.), *Handbook of public relations* (pp. 357–68). Newbury Park, CA: Sage Publications.

Rogers, E., & Storey, D. (1987). Communication Campaigns. In C. Berger & H. Chaffee (Eds.), *Reader in Public Opinion and Mass Communication* (3rd ed.) (pp. 419-445). New York, NY: Free Press.

Rossi, P., Lipsey, M., & Freeman, H. (2007). *Evaluation: a systematic approach.* Thousand Oaks CA: Sage Publications.

Schwarz, A., & Fritsch, A. (2014). Communicating on Behalf of Global Civil Society: Management and Coordination of Public Relations in International Nongovernmental Organizations. *Journal of Public Relations Research*, 26(2), 161-183.

Scott, R. W. (1995). *Institutions and Organizations*. Newbury Park, CA: Sage Publications.

Shulha, L., & Cousins, J. (1997). Evaluation Use: Theory, Research, and Practice Since 1986. *American Journal of Evaluation,* 18: 195-208.

Simmons, P. (1998). Learning to Live with NGOs. *Foreign Policy*, *112*, 82-96.

Sireau, N. (2009). *Make Poverty History: Political Communication in Action*. Hampshire, UK: Palgrave Macmillan.

Sriramesh, K. (2009). Globalisation and public relations: The past, present, and the future. *PRism* 6(2). Retrieved from:
 http://www.prismjournal.org/fileadmin/Praxis/Files/globalPR/SRIRAMESH.pdf

Starling, S. (2010). Monitoring and evaluating advocacy: lessons from Oxfam GB's Climate Change campaign, *Development in Practice*, *20* (2), 277-286.

Synder, L. (2001). How effective are mediated Health Campaigns? In Rice, R., & Atkin, C. (Eds.). *Public Communication Campaigns* (pp. 181-192). London: Sage publications.

Tkalac, A., & Pavicic, J. (2009). Nongovernmental organizations and international public relations. In K. Sriramesh & D, Verčič. (Eds.). *The global public relations handbook: Theory, research, and practice (*2nd ed., pp. 807-821) New York, NY: Routledge.

Trochim, W.M.K. (2009). Evaluation policy and evaluation practice. In W.M.K.

Trochim, M.M. Mark & L.J. Cooksy (Eds.), *Evaluation policy and evaluation practice*. *New Directions for Evaluation, 123,* 13–32.

Valente, T. (2001). Evaluating communication campaigns. In R. Rice & C. Atkins (Eds.), *Public communication campaigns* (pp. 105-124). Newbury Park, CA: Sage Publications.

van Ruler, B, Tkalac-Verčič, A., & Verčič D. (Eds.). (2008). *Public relations metrics.* New York, NY: Routledge.

Voils, C. I., Sandelowski, M., et al., (2008). Making Sense of Qualitative and Quantitative Findings in Mixed Research Synthesis Studies. *Field Methods, 20* (1), 3-25.

Walker, G. (1994). Communicating public relations research. *Journal of Public Relations Research,* 6(3), 141–61.

Watson, T. (1997). Measuring the success rate: Evaluating the PR process and PR programmes. In P. J. Kitchen (Ed.), *Public relations: Principles and practices* (pp. 283–99). Boston, IL: International Thomson Business Press.

Watson, T., & Noble, P. (2007). *Evaluating Public Relations: a best practice guide to public relations planning, research & evaluation* (2nd ed.). London: Kogan Page.

Weis, J., & Tschirhart, M. (1994). Public information campaigns as policy instruments. *Journal of Policy Analysis and Management*, *13*(1), 82-119.

Weiss, C.H. (1998). *Evaluation: Methods for Studying Programs and Policies* (2nd ed.). Newbury Park, CA: Sage Publications.

Welch, C. (2001). *NGOs and human rights: promise and performance*. Pennsylvania, PA: University of Pennsylvania Press.

White, H. (2008). Of Probits and Participation: The Use of Mixed Methods in Quantitative Impact Evaluation. *IDS Bulletin*, *39*(1), 98-109.

White, J. (2005). *Measurement and Evaluation: Moving the Debate Forward*. London: Chartered Institute of Public Relations.

Wholey, J. (1994). Assessing the feasibility and likely usefulness of evaluation. In J. Wholey, H. Hatry & K. Newcomer (Eds.), *Handbook of Practical Program Evaluation* (pp. 15-39). San Francisco, CA: Jossey-Bass.

Willetts, P. (2010). *Non-governmental organizations in world politics: the construction of global governance*. London: Routledge.

Williams, B., & Imam, I. (Eds.) (2007). *Systems Concepts in Evaluation: An Expert Anthology*. Point Reyes, CA: Edge Press/American Evaluation Society.

Xavier, R., Mehta, A., & Gregory, A. (2006). Evaluation in Use: The Practitioner View of Effective Evaluation, *Prism 4*(2). Retrieved from: http://www.prismjournal.org/fileadmin/Praxis/Files/Journal_Files/Evaluation_Issue/XAVIER_ET_AL_ARTICLE.pdf

Xavier, R., Patel, A., Johnston, K., Watson, T., & Simmons, P. (2005). Using evaluation techniques and performance claims to demonstrate public relations impact: An Australian perspective. *Public Relations Review, 31*(3), 417–24.

Zerfass, A. (2008). The corporate communication scorecard: A framework for managing and evaluating communication strategies. In van Ruler, B., Tkalac-Verčič, A., & Verčič, D. (Eds.), *Public relations metrics* (pp. 139-153). New York, NY: Routledge.

Interlude 2

Article 1 provided an overview of current practices of communication evaluation within IOs, with it found to be occurring in a minority of organisations. When evaluation was carried out, it was largely a post-activity action carried out by external consultants, using a narrow selection of methods and less than rigorous designs, although it focused more on outcomes compared to other sectors. The article concludes that it is both feasible and appropriate for IOs to adopt more diverse methods and rigorous designs, as evident in a minority of evaluations studied. However, it is suggested that the ability of organisations to do so is influenced by factors outside of the evaluation process such as the integration of evaluation within the communication function and the strength of the evaluation institution and its policies, in addition to other possible organisational and contextual factors.

While article 1 focused on the first component of the evaluation process, Methodology, article 2 focuses on the second component, Implementation. Through the evaluation of two communication campaigns of two IOs by this author, an analysis and reflection on the methodology used and its implementation is provided. This extends the proposals of article 1 as to what is feasible and appropriate, and allows a more in-depth exploration of the influence of both internal and external factors.

5.   Article 2 – Implementing communication evaluation methodologies for two international organisations

Abstract

Through the experience of evaluating two communication campaigns of the International Committee of the Red Cross and the Office of the High Commissioner for Human Rights, this article provides an analysis and reflection on challenges for communication evaluation methodology and its implementation leading to insights as to what is appropriate and feasible considering factors of influence, both internal and external. An appropriate response found was the pragmatic and adaptive approach adopted for the evaluations; an attempt to find the "middle ground" between the logical and linear evaluation process and the complex activities and contexts of the two campaigns. This produced broad findings on the campaigns and efficiency of their activities but compromised the range of methods that could be used and the consequent strength of the findings. This was tempered by the participative nature of the evaluations that built confidence in its methods and findings. Internal rather than external factors were found to be more important in influencing the methodology adopted and implemented: the campaign model; the level of integration of evaluation within the communication function; and the organisational settings. Although the experience showed that organisations could adopt appropriate evaluation methodologies, designs and methods, the feasibility to do so was reliant on their ability to counter and/or capitalise on the factors of influence, which in turn would facilitate communication evaluation that is both robust and of value to the organisations.

Keywords: communication, campaigns, evaluation, international organisations, non-profit communications, insider-outsider research, evaluation methodology

5.1. Introduction

Evaluation of communication activities has been named consistently as one of the top concerns of communication professionals globally (Gregory & White, 2008; Zerfass, Verčič, Verhoeven, Moreno & Tench, 2012) but paradoxically not even half reportedly undertake any evaluation (Fischer, 1995; Macnamara, 2006; Pohl & Vandeventer, 2001; Walker, 1994; Watson, 1997; Xavier, Patel, Johnston, Watson & Simmons, 2005). Reasons for this disparity include the accessibility of communication professionals to data; the impracticality and complexity of methodology required; the vagueness of communication programme design and planning; the lack of resources and know-how of evaluation; and the absence of an evaluation culture amongst communications professionals (Cutlip, Center & Broom, 2006; Fischer, 1995; Kelly, 2001; Macnamara, 2006; White, 2005).

Amongst international organisations (IOs) little is known as to how these organisations plan, manage and evaluate their communication activities (Lehmann, 1999; Manheim, 2011; Sireau, 2009; Tkalac & Pavicic, 2009). A recent review by this author of evaluation reports and guidelines of IOs from 1995-2010 found that evaluation was prevalent in only 13% of these organisations (compared to 30-50% in other sectors) and that evaluation was not integrated, adopted widely or rigorously applied. Found to influence the communication evaluation process was the strength of the evaluation institution and its policies, the level of integration of evaluation within the communication function and possibly other elements of the organisational and external contexts (O'Neil, 2013).

Given these findings, this author sought to further understand the challenges in implementing communication evaluation methodologies by collaborating with two IOs, the International Committee of the Red Cross (ICRC) and the Office of the High Commissioner for Human Rights (OHCHR) in the evaluation of two global communication campaigns. This article provides an analysis and reflection on carrying out the two evaluations and examines the internal and external factors that influenced

implementation. In doing so, this provides insights into appropriateness (what is suitable for the organisations) and feasibility (what in practice is possible to do) of evaluation methodology for communication activities of IOs. This article is the second in a series of articles as part of this author's PhD research.

Created through international or regional treaties, intergovernmental organisations (IGOs) undertake a variety of functions including cooperation, monitoring, dispute settlement or humanitarian intervention (Alvarez, 2006). Private organisations, international non-governmental organisations (INGOs) operate in multiple countries, normally with a voluntary aspect, whose members come together for a common purpose (Mingst, 2004), acting as advocates for specific policies, mobilising publics, monitoring or providing humanitarian assistance. Both IGOs and INGOs are considered as international non-state actors that form the "international public" or "international non-profit" sector (Charnovitz, 2006) and are referred to as international organisations (IOs) in this article.

Today, IOs are thought to be some of the most important influences on states, framing and reframing national, regional and international debates (Barnett & Finnemore, 1999; Elman and Elman, 2003). A key activity for these organisations is the collection and analysis of information – and its communications (Keck and Sikkink, 1998).   These organisations actively manage communication through programmes, projects and campaigns (Grunig, 1992). The way they communicate has massively changed in the past twenty years, with organisations now more global, strategic and interactive in their communication, even if research shows their potential is yet to be reached (Grunig, 2009; Sriramesh, 2009). Today, some of these organisations are household names in many countries: UNICEF, World Food Programme, Greenpeace and Oxfam amongst others (Kaldor, Moore & Selchow, 2012).

Communication campaigns are a distinct sub-set of broader communication programmes (Dozier, Grunig & Grunig, 2001) and the particular focus of this article. This sub-set is defined as an organised set of communication activities, directed at a particular audience usually within a defined period of time to achieve specific outcomes (Rogers and Storey,

1987). Coffman (2002) distinguished between two types of campaigns: 1) public information campaigns that aim to inform and influence the behaviours of individuals and 2) public will campaigns that aim to mobilise public action for policy change. IOs use campaigns for a combination of both goals: UNICEF campaigns on the importance of child education targeting both parents and policy makers (Chesterton, 2004); Oxfam campaigns on the risks of climate change seeking to mobilise publics to put pressure on their governments and also seeks to influence behaviour of individuals (Cugelman & Otero, 2010); and the International Organisation for Migration (IOM) campaigns on developing positive attitudes of publics towards refugees and asylum seekers (IOM, 2002).

5.2. Challenges in communication evaluation and its implementation

For the past seventy years, numerous communication campaigns have been evaluated and their results published, although most campaigns evaluated have been described as "small scale campaigns" (p.40, Salmon, 1989) in terms of ambition and geographical coverage and mainly in the public health area overseen by government agencies or local NGOs (Coffman, 2002). The commonly accepted characteristics of campaigns – achieving specific outcomes within a set time period in theory make them more suitable to evaluation than communication programmes that typically have broader objectives and run continuously or with rolling time periods (Wilson & Ogden, 2008).

Historically, campaign evaluation has focused on determining to what extent campaigns could influence knowledge, attitudes and behaviour. An often cited research study titled "Some reasons why information campaigns fail" (Hyman & Sheatsley, 1947) stated some of the limitations of campaigns in inducing change. Some forty years later, Salmon (1989) argued that "campaigns indisputably are capable of inducing effects" (p.40) although cautioning that success was a subjective determination based on the interpretation of evaluation data. Today, campaign evaluation still has a focus on determining the level of effects seen but also emphasised is the purpose of evaluation for learning and continuous improvement of organisations reflecting a general shift of evaluation in this direction

(Communications Consortium Media Centre [CCMC], 2004; Michaelson & Macleod, 2007).

A challenge seen in communication evaluation is the lack of diversity of methodological foundations and consequent methods used. The available methods for campaign evaluation have been considered as "vastly deficient" (p. 12, Coffman, 2002), reflecting the lack of methodological diversity and inability to use appropriate methods in communication evaluation in general (Cutler, 2004; Dozier, 1990; Jelen, 2008). This lack of diversity in methods used has also been found in IOs (O'Neil, 2013). Some authors also point out that epistemological and methodological foundations have not been addressed nor have the paradigms debates found in the broader evaluation field (Datta, 1994; van Ruler, Tkalac-Verčič & Verčič, 2008).

Criticism has also been made of the limited and isolated focus of communication evaluation. The theories and models developed for communication evaluation have largely a narrow emphasis on evaluating the communications programme or campaign with limited consideration given to the interrelations between the evaluation process and its organisational setting and context (Lennie & Tacchi, 2013; Likely & Watson, 2013).

This narrow focus is also reflected in the distinct preference for positivist linear models of past campaign evaluations. Quasi-experimental designs with pre-tests, post-tests and control groups are dominant (Valente, 2001) although this has been increasingly challenged as not always appropriate or feasible given complexities of communication activities, contexts and audiences (Broom & Dozier, 1990; Coffman, 2002; Flay & Cook, 1981; Lennie & Tacchi, 2013).

The logical and linear processes used have been questioned also on the basis that communication activities will not necessarily be implemented as planned, producing unpredictable results and consequently change and adapt to the given context and environment (Lennie & Tacchi, 2013; Manheim, 2011). Woolcock (2009) considers that this can lead to a mismatch between the planned and actual "impact trajectory" that

inevitably leads to errors in evaluation. The example of an AIDS awareness campaign is given, where planning and evaluation may be based on the assumed trajectory of a gradual raise in awareness but in reality may be closer to a horizontal 'S' curve trajectory, with awareness peaking and falling in waves.

Rogers (2008) found that evaluation methodology in general has tended to underestimate the challenges of evaluating complicated activities (multi-level and multi-site) with emergent (rather than pre-defined) outcomes in complex environments.   IOs have been found to not yet fully recognise the complexity of the global contexts of their communication activities and consequent evaluation approaches required (Lennie & Tacchi, 2013).

To cope with issues of context and complexity, alternative approaches have been developed such as non-linear models and adapted evaluation designs, drawing from systems theory (Williams & Imam, 2007). Bamberger, Rugh & Mabry (2011) have studied these issues and offer practical guidance on conducting evaluations at an acceptable level of rigour and validity when faced with organisational and resource constraint through adopting an integrated approach combining a wide range of evaluation tools.

Patton (2011) cautioned that evaluation in its implementation finds itself in the "muddled middle" (p.177), pressured simultaneously from the "top-down" that demands standardized procedures and from the "bottom-up" of contextual factors and programme adaptation that exist in reality. He advocates "methodological appropriateness" (p. 290) through the use of multiple and diverse methods given that uncertainty and non-linearity are a given in most contexts.   Schwandt (2003) described this as the "rough ground" (p. 355) of reality where theories and concepts clash in their implementation.

Through the experience of evaluating two communication campaigns, this article considers these challenges of methodology and implementation directly confronting the issues faced at the "rough ground" by the "muddled middle", a perspective that has been

described in the literature but little studied in-depth through empirical studies such as this one (Bamberger et al., 2011; Patton, 2011, Weiss 1998).

5.3. Methods: three elements

Three methodological elements were used for this article. The first element was the methodology that has been used to evaluate the communication campaigns. Both quantitative and qualitative methods were used inspired by a nested analysis mixed methods strategy (Lieberman, 2005) guided by an overall adaptive and pragmatic approach. This methodology endeavours to counter the complexity of contexts and consequent programmes, by breaking down the latter into smaller elements and matching it to different evaluation methods and approaches that is then "nested" in an overall design. With these two campaign evaluations, quantitative methods (e.g. surveys with target publics) and qualitative methods (e.g. interviews with campaign staff and partners) were used for different parts of the evaluation and nested in a theory-based design that then validated the extent to which the two campaigns achieved their set objectives.

The second methodological element used by this author was to reflect on the experience of carrying out the two campaigns and its broader implications for this field. For this purpose, the author drew upon the "insider-researcher" approach (Mason, 2002; Radnor, 2001) although the role was closer to what has been labelled as the "insider-outsider" approach; that is the "space between" insiders and outsiders (Dwyer & Buckle, 2009; Kerstetter, 2012). The author was an outsider as an external consultant but at the same time an insider given that the research was done closely with both organisations for some 18 (OHCHR) and 28 (ICRC) months respectively. In this approach, there is also a recognition that the researcher's identity can change during the research process, depending upon factors such as where and when the research is being conducted, the participants involved and the context (Dwyer & Buckle, 2009).

"Insider-outsider" research can be considered as a type of action research although this article does not purport to be an action research piece, given that a defining element of

action research is the internal cycles of action and change (Reason & Bradbury, 2001). However, action research provides some background for the methodology adopted given that its roots lay largely in the work of Kurt Lewin (1948) whose original research setting was that of an external consultant working alongside an organisation in a collaborative process, similar to the role taken by this author. The author used several tools drawn from action research to record the interaction with the organisations; a personal log of field notes describing the main contact with the organisations and a record sheet recording the milestones of the evaluations as they progressed (McNiff & Whitehead, 2011).

The role played by this author was also more nuanced than that of an external consultant in that it was that of an evaluator, a role that itself is open to different interpretations. The variations seen for this role reflect those seen in methodological approaches and purposes of evaluation (Luo, 2010). Campbell (1984) and Scriven (1986) saw evaluators as "methodologists" and "judges" that worked independently from organisations to assess the outcome effects of programmes. As "educators" (Weiss, 1998; Wholey, 2010) and "facilitators" (Patton, 2000; Stake, 1980) these authors saw evaluators focus on educating and involving the organisations and stakeholders in the evaluation process and ensuring that they produced findings for a programme's improvement. Rossi, Lipsey and Freeman (2007) saw multiple roles which would change based on the different stages of the evaluation, similar to the changing role of the "Insider-outsider" researcher as described above.

Discussion on the evaluator's role is largely absent from the campaign and communication evaluation literature, as is any broader reflection on the actual experience and implementation of evaluation. Where concern has been expressed is the tendency towards viewing evaluation as a distinct action carried out by an external "evaluator" and the lack of emphasis in the literature on integrating evaluation within the communication function (Broom & Dozier, 1990; Watson & Noble, 2007). This tendency of evaluation as an action carried out by external consultants distinct from the communication function was also found in this author's study of IOs (O'Neil, 2013). Mendelsohn (1973) believed

that external evaluators and the reports they produced are viewed with "a great deal of suspicion and distrust" (p. 51) by communication professionals due to this external role and its lack of integration within the communication function.

Reflecting on a researcher's own experience requires interpretation and guidance. MacIntyre as quoted in Mason (2002), identified three levels of reflection: (i) technical: the description of specific actions, (ii) practical: questioning the assumptions and consequences which link actions and (iii) critical: considering the wider societal and institutional issues that may limit the efficacy of these actions. This author utilises these levels as a guide with technical reflection largely used in describing the evaluation process and steps carried out. Practical reflection is used when considering the components of the evaluation with critical reflection used in the discussion and conclusions when considering factors that influenced the evaluation process.

The third methodological element is a conceptual framework for communication evaluation for IOs as seen in Figure I. Developed by this author, this framework sets out the evaluation process in relation to internal and external factors that are considered to have an influence on an organisation's ability to carry out communication evaluation. The framework draws from systems theory, organisational behaviour and existing empirical and theoretical research in the communication and evaluation fields (Bamberger et al., 2011; Christie & Fierro, 2012; Dubin, 1976; Fitzpatrick, 2012; Højlund, 2014a; Manheim, 2011; Meredith, 1992; Scott, 1995; Trochim, 2009; Watson & Noble, 2007).

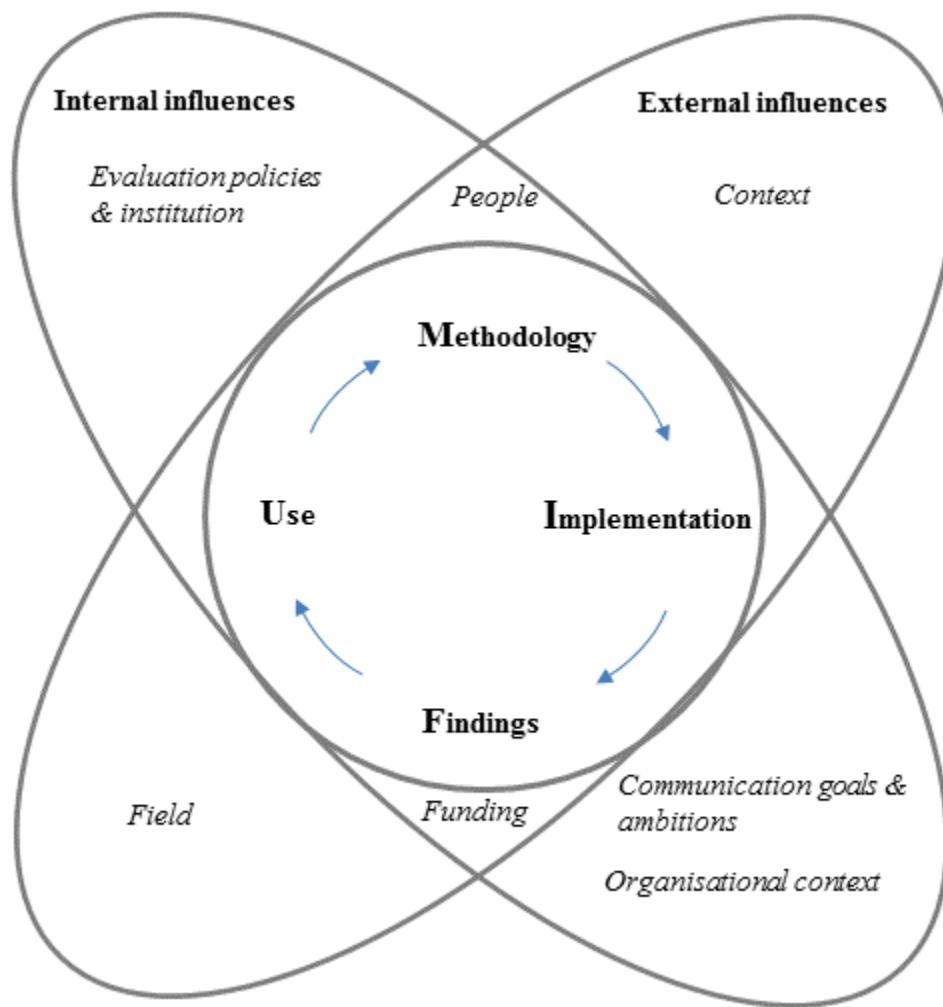*Figure I: MIFU conceptual framework for communication evaluation for IOs*

The framework highlights the interaction between the four components of the communication evaluation process (Methodology, Implementation, Findings and Use) and the possible influence from factors outside but interlinked to this process. These factors are described in Table I below, based on the above-mentioned references.

Table I: Internal and external factors of the conceptual framework	
Evaluation policies and institutions	Internal to the organisation, evaluation policies, guidelines and direction as managed by a central evaluation unit; the main components of the larger evaluation system.
Organisational context	Internal contextual elements such as structure, culture and strategy.
Communication goals and ambitions	The scope of the communication activities being evaluated, such as the level of effects being sought and the implementation models used.
Context	The external setting within which an organisation carries out its communication activities.
Field	The community of organisations and bodies that have a common meaning system (e.g. peer organisations, academia and industry associations). Two fields are relevant for this study: evaluation and communication evaluation.
Funding	The financial resources available to communication evaluation.
People	The competencies of the human resources implicated in communication evaluation (e.g. communication staff, evaluation staff and external evaluation consultants).

5.4. Two communication campaigns

The two organisations and their campaigns are now further described. These two organisations were selected as firstly, they were both within the definition of what are considered as IOs, secondly as they both had organised and visible communication activities and thirdly, they were willing to collaborate for this research. The ICRC is the founding body of the International Red Cross and Red Crescent Movement. The mandate of the ICRC focuses on protecting and assisting the victims of armed conflict. The ICRC operates in some 80 countries and employs 12,000 persons worldwide (International Committee of the Red Cross, 2010). The OHCHR is a component of the United Nations system. The mandate of OHCHR focuses on promoting and protecting human rights. OHCHR operates in some 60 countries and employs 1,300 persons worldwide (Office of

the High Commissioner for Human Rights, 2012). Both organisations have their headquarters in Geneva, Switzerland. From December 2007 to December 2008, the OHCHR conducted a global campaign on the 60th anniversary of the Universal Declaration of Human Rights (UDHR). This campaign ran under the title "Dignity and justice for all of us" and its goal was to increase knowledge and awareness on human rights and empower people to realise their rights. The campaign was conducted with partners from within and outside the UN system.

 From January 2009 to December 2009, the ICRC conducted a global campaign on the 60th anniversary of the Geneva Conventions. The campaign ran under the title "Our world. Your move." and its goal was to focus the world's attention on the value of humanity and the humanitarian gesture. The campaign was conducted with partners from within the International Red Cross and Red Crescent Movement. Table II summarises the main characteristics of the two campaigns (further detailed in annex two of this thesis):

Table II: Main characteristics of the two campaigns		
	ICRC	OHCHR
Title	Our world. Your move	Dignity and justice for all of us
Time period	January 2009 – December 2009	December 2007 – December 2008
Goal	To focus the world's attention on the value of humanity and the humanitarian gesture	To protect and empower people to realise their human rights
Intended effects	Changes to knowledge, attitude & behaviour of individuals	Changes to knowledge, attitude & behaviour of individuals Changes to policies, activities & practices of institutions
Message focus	60th anniversary of the Geneva Conventions 150th anniversary - Solferino battle	60th anniversary of the Universal Declaration of Human Rights (UDHR)
Activities	Media relations and events, online activities, support to local activities, exhibitions, research on contexts and consequent material	Media relations and events, artistic projects, grassroots mobilisations, support to local activities
Campaign partners	Red Cross Red Crescent National Societies and their International Federation	National Human Rights Institutions, national NGOs, government entities, education institutions
Geographical coverage	Global	Global
Type of campaign	Public information	Public information, public will

The campaigns were similar in many respects; they were both global; they both focused on anniversaries of international treaties in the human rights / humanitarian field; they both used a partnership strategy for campaigning and they both used a similar range of activities. These similarities facilitated the analysis for this research as detailed below.

5.5. The evaluation process

To initiate the evaluations, in mid-2008, the author made contact with both organisations and had an initial discussion with the communication staff about the possibility of evaluating a current or planned global campaign.

At this time, OHCHR had launched their global campaign (as described in Table II) some five months earlier and were occupied with the implementation of their campaign activities. OHCHR had a communication unit with five staff that were working on the campaign in addition to other communication activities. No staff were permanently dedicated to the campaign although it occupied some 80% of the time of two staff. The ICRC was preparing to launch their global campaign (as described in Table II) seven months later and had a three person team dedicated to the campaign and part-time support from some 10 staff spread across the various communication sub-function, e.g. media, social media and audio-visual production (G. O'Neil, field notes, 29 May 2008 & 25 June 2008). In both organisations, the main contact for this author ("the evaluator") was the person responsible for the overall management of the campaign ("campaign manager").

The evaluator adopted a participatory approach and aimed to include the communication staff in all steps of the evaluation, working with and through the respective campaign managers (Cousins & Chouinard, 2012). In this regard, the evaluator's own experiences and preferences came into play and influenced the various elements of both the methodology and implementation steps of the evaluation, as detailed in the central circle of Figure I. The role of the individual evaluator or team in shaping the nature and approach of the evaluation has long been recognised in the literature (Patton, 2008; Weiss, 1998).

For example, both organisations were guided by the author in the procedures to follow in managing the evaluation, such as the level of consultation needed, input into findings and their validation. At the same time, both evaluations were initiated by the communication functions directly and there was limited input from the central evaluation policy and institution. In discussions with the campaign managers it was understood that the evaluations had to respect the broad principles of evaluation established in the organisations and mostly operationalised by results-based management (RBM) systems in place in all functions (including communication), such as the focus on the outcome level for monitoring and evaluation, the establishing of links between activities, objectives and outcomes and the emphasis on organisational learning. At the time of the campaign evaluation, OHCHR had just introduced an organisation-wide RBM system and ICRC some 10 years earlier (G. O'Neil, field notes, 26 September 2008 & 27 October 2008).

The process for carrying out the evaluations was agreed with both organisations and followed six main steps as recommended for campaign evaluation (Coffman, 2002; Dubey & Bardhan, 1981). These steps fitted within four components of the evaluation process found in the central circle of Figure I: Methodology, Implementation, Findings and Use.

The first component, Methodology, contained steps one to three. These were essentially the preparatory steps for the evaluation. Step one involved defining with the two organisations the desired outputs and outcomes of the campaigns linked to the set objectives. The importance of defining objectives for the communication activities to be evaluated has been emphasised in the evaluation literature (Lindenmann, 2003; Shaw, Mark & Greene, 2006; Weiss, 1998). Together with the campaign managers, based on the campaign objectives and goals, desired outputs/outcomes of each campaign were defined, as detailed in Table III. To facilitate the evaluation, indicators (measurable points indicating progress), were then determined for the objectives: 17 indicators for the seven outputs/outcomes of the OHCHR campaign and 18 indicators for the 10 outputs/outcomes of the ICRC campaign. This was carried out through an iterative

process with the organisations, with this author drafting indicators for outputs/outcomes, discussing and modifying them in consultation with the campaign managers.

Table III: Desired outputs/outcomes for two campaigns	
ICRC	*OHCHR*
1. Strengthened communication network between ICRC, International Federation & National Red Cross/Crescent Societies (NS) 2. Developed partnerships for ICRC & International Federation within the Movement 3. Increased capacity for ICRC & International Federation to use new media & social networking 4. Increased awareness of today's most pressing humanitarian challenges (conflict & climate change) 5. Increased awareness of needs, vulnerabilities & expectations of beneficiaries 6. Increased association of ICRC, International Federation & NS as key actors for today's humanitarian challenges 7. Motivated people to undertake a humanitarian gesture 8. Motivated people to undertake voluntary work 9. Motivated people to undertake an online action of support 10. Raised funds for the ICRC, International Federation & NS	1. UN system actively participated in the campaign 2. Engaged multiple stakeholders in the campaign at the country level & globally 3. Increased awareness of Universal Declaration of Human Rights and human rights in general amongst rights holders 4. Increased awareness of human rights relating to detention 5. Stimulated debate, spurred action and reaffirmed commitment of governments, civil society, educational, cultural and human rights institutions to values and principles of UDHR 6. Helped bridge gaps in HR implementation at the national level 7. Garnered further political and financial support for OHCHR

Step two involved constructing and confirming the "theory of change" for each campaign in a collaborative process with both organisations. Theories of change describe the activities of the campaign in connection to what is trying to be achieved (outcomes) and shows the pathways and factors through which change is expected to occur (Coffman, 2003). In this regard, the theory of change is not a theory that guides the evaluation

process but more so a tool to understand how the campaigns were intended to work and what were the theoretical underpinnings to bring about change (Funnell & Rogers, 2011).

Step three involved designing the research methods to evaluate the campaigns. The approach taken was to work with multiple types of qualitative and quantitative methodologies and tools framed by a nested analysis mixed methods strategy. Discussions were held with both organisations to determine the feasibility of the methods and the sampling strategies proposed. This was necessary given the range of outputs and outcomes to be evaluated and a reflection of the current tendencies in communication evaluation, that is, to use multiple methods instead of a single method (CCMC, 2004; Gregory, Morgan & Kelly, 2005; Lindenmann, 2003). For both campaigns, the methods used were content analysis of media coverage, semi-structured interviews with campaign partners and staff, surveys with campaign partners and select audiences and analysis of monitoring data. The methods were then matched to objectives and indicators described above.

At this stage, an evaluation framework was created for both evaluations detailing the campaign outputs/outcomes, indicators, theories of changes, research methods, sampling strategies and a timeframe. The frameworks were discussed and validated with the campaign managers which then guided the implementation of the evaluations (G. O'Neil, record sheet, 4 December 2008 & 19 January 2009).

Implementation involved steps four and five. Step four involved the deployment of the research tools in collaboration with the two organisations. The OHCHR evaluation was a post-only design and therefore data collection was undertaken after the campaign was completed (early 2009). The ICRC evaluation was a pre-post design with data collection undertaken prior to the campaign (late 2008), during the campaign and after its completion (early 2010). The data collection was carried out jointly with the organisations. For example, both organisations carried out monitoring activities, such as monitoring media coverage and tracking participation of partners in the campaigns that were used for the evaluation. The author carried out other data collection independently

such as surveys of audiences and interviews with campaign staff and partners while having validated by the organisations any tools used, such as interview guides, content analysis frames and survey questions. During the data collection, the evaluator was in contact with the campaign managers for general coordination purposes and to discuss the results emerging. Once the data was collected, the evaluator carried out the analysis (step five) using appropriate analytical approaches. The theories of change were then re-examined in light of the findings to determine if progress was achieved as desired. The following diagram illustrates the interaction between the evaluator and the organisations in designing the research methods, data collection, data analysis and presentation steps.
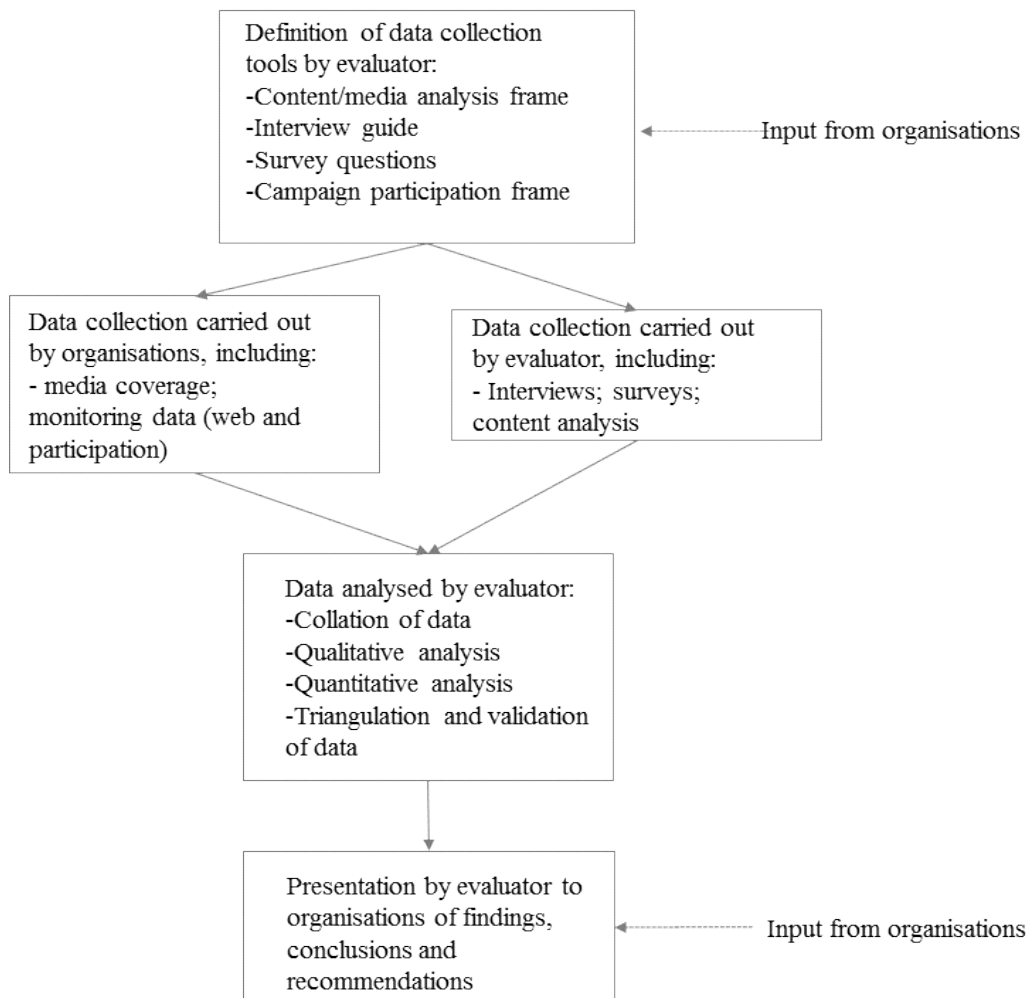
Figure II:Interactions between evaluator and organisations in data collection and analysis

The above-mentioned outputs and outcomes were categorised into seven levels of anticipated effects split between outputs and outcomes (Macnamara, 2005; Rogers & Storey, 1987). The following table illustrates the data collection in relation to the levels of effect and the evaluation design of the campaign evaluations.

Table IV: Levels and data collected for two campaigns by research design				
	ICRC		OHCHR	
	Pre-post data	Post-only data	Pre-post data	Post-only data
Output				
1. Level of participation	X			X
2. Change to organisational capacity		X	N/A	
3. Visibility in the media	X		X	
Outcome				
4. Change to knowledge – awareness	X			X
5. Change to attitude	X			X
6. Change to behaviour – individual action		X	N/A	
7. Change to behaviour – organisational action	N/A		X	

The Findings component contained step six. This step involved discussing, presenting and disseminating the evaluation findings to and within the two organisations. For both organisations, the results were shared in draft report format for comments from the campaign staff and management (G. O'Neil, record sheet, 7 July 2009 & 6 June 2010). Once comments were received and appropriate modifications undertaken, the reports were then finalised (G. O'Neil, record sheet, 4 September 2009 & 30 August 2010). In both organisations, the evaluation findings, conclusions and recommendations were then presented by the evaluator to a broader group of management, communication and other interested staff in a findings workshop (G. O'Neil, record sheet, 2 October 2009 & 14 September 2010).

At this stage, the evaluation process involving the evaluator was concluded. For the final component of the evaluation process, "Use", the responsibility was handed over to the organisations. Based on the discussion in the findings workshops, it was anticipated that the evaluations would be used within the organisations as part of their learning processes and for designing future campaigns (G. O'Neil, field notes, 2 October 2009 & 14 September 2010). The usage of the evaluation findings within the organisations and the changes they induce or not, are an important area of the evaluation process and an indication of the overall value of evaluation (Patton, 2008). Findings are intended to feed into future evaluations and programming, thus the circular notion of the evaluation process as seen in Figure I and represented in other evaluation models, but not all (Pawson and Tilley, 1997; Weiss 1998; Wholey, 2004). However, the aspect of evaluation use is not considered in this article, but more so the focus of the third article of this PhD research.

The interactions between the evaluator and the two organisations and the main milestones of the evaluations throughout the evaluation process are mapped out (in relative time) in Figure III. The upper tier shows the interactions and milestones with the ICRC and the lower tier with OHCHR. The figure shows a period of 29 months with the campaign duration of 12 months at its centre (both campaigns ran for 12 months).

A distinction between the two evaluations seen in this mapping is the length of time; the OHCHR evaluation was carried out over 18 months whereas the ICRC was 28 months. This was due to the nature of the evaluation design, as the ICRC was pre-post and OHCHR post-only. This also influenced the number of interactions of organisations; 15 with OHCHR and 18 with the ICRC. Interactions mapped included physical meetings with the campaign managers and their staff and do not include email exchanges, telephone calls or data collection, i.e. interviews with staff or partners.

The intensity of the collaboration differed between the two evaluations. With the OHCHR evaluation, there was more interaction between the initial discussion and finalising the evaluation framework than the ICRC; eight compared to three interactions.

On the contrary, there was more interaction during the data collection period (between the two milestones "Framework finalised" and "Draft report delivered") with the ICRC than with the OHCHR; eleven compared to six interactions.

For the difference in the initial phase, discussions on the framework and the evaluation were more intense with the OHCHR, given that it was the first time the organisation was carrying out an evaluation of their communication activities and with the evaluator, so more explanation and discussion was needed than with the ICRC (G. O'Neil, field notes, 29 May 2008). With the ICRC, they had previously carried out evaluations of communication activities (although not of a global campaign) and the evaluator had previously collaborated with them on several smaller evaluation projects (in addition to being a former staff member 10 years earlier). Therefore, less discussion and explanation was required than with OHCHR (G. O'Neil, field notes, 25 June 2008).

In relation to the data collection phase, as the ICRC evaluation was a pre-post design, the evaluator was active in data collection during the campaign which required coordination. In addition, the campaign staff were interested to learn of any findings as they emerged in order to adapt their approach and tactics, which was not possible for the OHCHR evaluation as it was a post-only design.

Several similarities between the organisations can also be seen in the evaluation processes. For example, time periods between the milestones of "Initial discussions" and "Framework finalised" was similar; six months for the ICRC and eight for OHCHR. Between the "Draft report delivered", "Draft report finalised" and "Findings workshop", the time period was two months and one month respectively for both organisations.
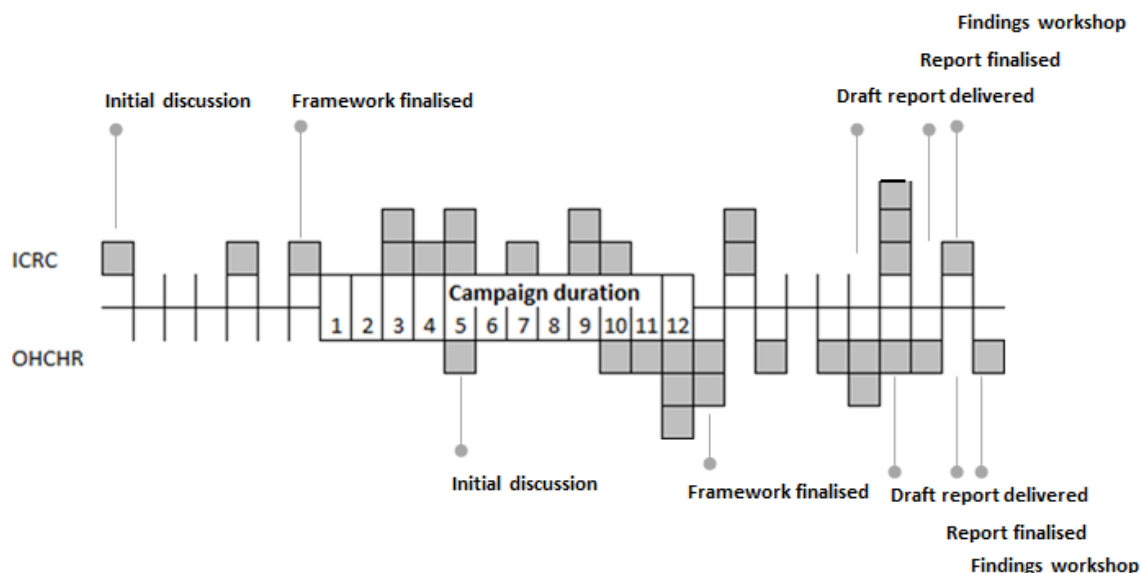
Figure III: Interactions and milestones of the two campaign evaluations

> *Notes for Figure III*
>
> One greyed square is equal to one interaction (physical meeting) with the organisation; the unit of time is months with the vertical lines separating each month.

5.6. Results: analysis and reflection on the evaluations

5.6.1. Analysis and reflection on the Methodology component

The experience of working with the two organisations within the Methodology component are now analysed and reflected upon in this section.

The starting point for the evaluations was the identification of the campaign objectives and matching them to desired outputs/outcomes as described above. The ICRC had already determined broad objectives for their campaign when the initial discussions on the evaluation were held (G. O'Neil, field notes, 25 June 2008). The OHCHR campaign was five months into its implementation when initial discussions were held. Their campaign was based on broad goals that were then broken down into outputs/outcomes by the evaluator and the campaign manager as illustrated in Table III (G. O'Neil, field notes, 29 May 2008). For both campaigns, the broad nature of the objectives for the campaigns led to issues of evaluability, that is, the extent to which an activity can be

evaluated, which has been previously raised in the literature concerning communication and campaign objectives (Hon, 1998). An examination of the desired output/outcomes (table III) illustrates that the ambitions of the campaigns were to target globally dispersed audiences across different contexts with various desired effects sought. This in turn meant the creation of many indicators and the use of multiple methods as part of the evaluation, bringing with it various challenges, as described later in this article. Although there has been a move towards the use of multiple indicators in communication evaluation there has also been a recognition of the complexity it implies for evaluation and organisations (Gregory & White 2008; Hon, 1998; Salmon & Murray-Johnson, 2001).

Both organisations sought a focus on "outcome" level over the "output" level of effects as seen in Table III. This was on the basis that evaluation at the outcome level is of more significance and value to organisations (Lindenmann, 2003; Rogers and Storey, 1987) and was also encouraged by this author in discussions with the campaign managers and reflected in the consequent evaluation frameworks (G. O'Neil, field notes, 4 December 2008 & 19 January 2009). As mentioned above, the focus on outcomes was also reinforced by the prevalent evaluation policies of the organisations and reinforced by its systems, notably the RBM systems in place.

The use of multiple outcomes matched the global campaigning model adopted by both campaigns. This was distinct from a national or local campaign model that would often have one key objective with a limited number of target audiences within a set geographic region and specific effects sought, e.g. influencing commuter driving in one city, Atlanta, USA (Henry & Gordan, 2003) or influencing attitudes on aids amongst adolescent women in one country, Brazil (Porto, 2007). In determining the output/outcomes and consequent indicators, discussions were held with the campaign managers on the possibility of narrowing the output/outcomes to be evaluated, for example to examine more in-depth several objectives and their consequent activities. However, for both organisations, this was seen as an unacceptable compromise as they desired to have as comprehensive an evaluation as possible of the campaigns, that is, to assess all possible outcomes (G. O'Neil, field notes, 29 May 2008 & 25 June 2008). The consequences of

this approach were seen in the implementation component, where challenges were faced to fully evaluate all outputs and outcomes sufficiently as described below.

Concerning the development of the theories of change, for the ICRC this was developed prior to the commencement of the campaign. In the discussions with the campaign manager and her staff it helped in clarifying certain assumptions about how the campaign would bring about change, for example which activities were seen as contributing to which outcomes (G. O'Neil, field notes, 28 November 2008).   With the OHCHR, the theory of change was completed after the campaign had concluded (January 2009) and was useful in re-constructing with the campaign manager how change was supposed to occur but did not have any input into adjusting the campaign planning or activities (G. O'Neil, field notes, 28 January 2009). Given the broad scope of both campaigns, it emerged in the analysis step that the theories of change created did not address all the assumptions of the campaigns, notably the path from activities to short-term outcomes and then to long-term outcomes. This is an issue already evident in using theories of change in programme design and consequent evaluation (Vogel, 2012).

The selection of the evaluation designs guided the methods used. A pre-post design (without control group) was used for the ICRC campaign and a post-only design for the OHCHR campaign.   Some pre-post data could be re-constructed for the OHCHR campaign such as the level of media coverage and changes to organisational behaviour (in this case, the ratification or signature of human rights treaties) as seen in Table IV. With the ICRC campaign, a pre-post design was used which enabled some additional pre-post measurement to be carried out, such as a panel study on awareness levels. The possibility of carrying out a pre-post design with a control group was considered for the ICRC campaign but discounted given that the campaign sought maximum exposure to messaging and it was considered difficult to find a comparable audience that would have no potential exposure to the campaign (G. O'Neil, field notes, 28 November, 2008). This is an obstacle highlighted previously in the use of control groups within evaluation design for campaign evaluation (Flay & Cook, 1981).

The selection of the design also influenced many other aspects of the evaluations: the level of interaction with the organisations; the length of time of the evaluations; the ability of the evaluations to influence campaign objectives; and the ability to integrate emerging findings during the campaign. In this regard, the use of a pre-post design had obvious advantages over that of a post-only design.

Evaluation methods were selected based on a matching against the set indicators for each objective. The identification and selection of appropriate methods was discussed jointly with both organisations, guided by the evaluation designs and the resource limitations of the organisations and the evaluator, notably in staff and budgets available for the evaluations. Where possible, existing resources were used if they did not overly compromise the evaluation methodology. For example, to collect data on media visibility generated by the campaigns, both organisations had existing contracts with third-party companies to provide them with such data. This data source was used but meant that the evaluator was limited in influencing the corpus of media sources that made up the media databases; this was considered a compromise that was acceptable for the evaluations (G. O'Neil, field notes, 25 November 2008 & 28 November, 2008).

In discussions, with the organisations, various methods were considered, debated and their use validated or not. For example, with OHCHR, consideration was given as to whether it would be possible to place evaluation questions within global-level omnibus surveys, but this was discounted due to the envisaged costs and the difficulties seen in reaching relevant audiences (G. O'Neil, field notes, 25 November, 2008). At the same time, several methods were used in the implementation component which were not foreseen in this preparatory stage. For example, with the ICRC, a rolling survey incorporating questions on the campaign was conducted of Red Cross and Red Crescent National Societies by their peak body, the International Federation and these results were then integrated within the evaluation findings (G. O'Neil, field notes, 7 July 2010).

In selecting the evaluation methods, there was an awareness amongst the campaign managers and their staff that challenges would be seen in their implementation. These

challenges included the difficult access to audiences given the global nature of the campaigns and the need to rely on partners to collect some data. These challenges were discussed and it was decided to pursue all methods knowing that the risk existed that not all methods would be successful in the collection of sufficient and/or relevant data (G. O'Neil, field notes, 25 November, 2008 & 28 November, 2008).

Procedures such as the setting of the main milestones, planning and responsibilities for the evaluation illustrated that the same approach was possible to apply to both campaigns. The evaluator in agreement with the campaign managers could guide the evaluation process and consequent procedures, as the communication functions did not have set or fixed procedures for carrying out evaluation (G. O'Neil, field notes, 25 November, 2008 & 28 November, 2008).

 At a conceptual level, the methodology adopted was guided by the epistemological orientations of both the organisations and the evaluator. These aspects were not dictated by a preferred paradigm or approach of either organisation or the evaluator, for example in imposing certain evaluations designs, methods or principles. More so, through discussions between the campaign manager, staff and the evaluator, a preferred set of methodological approaches emerged, which considered the factors such as organisational setting (e.g. the flexibility given to the campaign managers), campaign model (e.g. the multiple objectives and activities, the set-up for campaigning) and context (e.g. the dispersed and diverse audiences). These approaches were consolidated in a nested analysis mixed methods strategy for the evaluations drawing from a theory-based model. This was combined with a participative and collaborative approach of the evaluator that was fitted within an overall linear structure atypical to evaluation. Therefore, any predispositions from epistemological orientations were countered by this combination of approaches resulting in an eclectic and flexible conceptual frame.

This combination of approaches was also reflected in the role the evaluator played during this methodology component of the evaluation. As a "facilitator", this role involved several aspects: discussing the proposed methodological aspects with the campaign

managers and staff; weighing up the risks and opportunities in the options available; taking decisions jointly with the campaign managers; and documenting and seeking validation of the evaluation planning, design and methods. This collaborative approach was adopted to counter the envisaged resistance of the communication staff to external evaluation documented by Mendelsohn (1973) and based on the link that has been found between staff participation and consequent confidence in methods used, ownership of the evaluation findings and their eventual usage (Ayers, 1987; Braverman & Arnold, 2008; Cousins, 1995; Greene, 1988; Shulha & Cousins, 1997). The evaluator was leading the process of proposing and advocating for appropriate methods and approaches and thus encouraging decisions that would reinforce confidence in the evaluation while being aware of the resource limitations and other influences. This has been labelled as "the evaluator's balancing act" (p. 71) between advocating evaluation rigour but recognising the constraints of the particular organisational settings (Braverman & Arnold, 2008).

Overall, the analysis and reflections of the Methodology component describe a scenario where methodological choices were largely shaped through a joint view of evaluation of the organisations and the evaluator, considering their organisational settings, context and campaign model. The integration or consideration of evaluation within the campaign planning emerged as a key influence on the evaluation design, with the advantages seen of a pre-post design used for the ICRC evaluation. The influence of the organisation's evaluation institution and policies was neither formal nor prescribed but it was expected that the evaluations would align with its principles, mostly visible through the RBM systems. The global campaign model of both campaigns set a broad range of objectives to be evaluated. A consequence was that this influenced the selection of methods with some possible limitations foreseen in implementing these methods, such as access to audiences and reliance on partners to collect data. The set of methodological approaches adopted was an attempt to find the "middle ground" between the logical and linear process and the complex activities and contexts of the two campaigns.

5.6.2. Analysis and reflection on the Implementation component

The experience of working with the two organisations within the Implementation component are now analysed and reflected upon in this section.

The first key step in this component was the data collection which is discussed by the level of intended effects: output level; level of participation, organisational capacity and media visibility and outcome level; changes to knowledge and attitude; changes to individual behaviour; and changes to organisations (policies, practices and activities of institutions).

Measuring outputs was not dominant in the evaluations, as determined in the Methodology component. A common occurrence within communication evaluation has been to only consider media coverage (an output measure) in evaluating a campaign or activity which has resulted in the 'level substitution' phenomena where output level measures are substituted for more significant outcome level measures (Cutlip, Centre and Broom, 2006; Grunig, 2008). This limitation was recognised in these campaign evaluations and reflected in the data collection.

Both campaigns had desired outputs that focused on the level of participation and strengthening the network of partners involved in the campaign (ICRC output/outcomes 1 & 2; OHCHR output/outcomes 1 & 2). This output was particular to international campaigning where the organisation itself may not be always communicating directly with audiences it seeks to influence but rather through influencing partners and networks to do so (Manheim, 2011). Assessing the participation of partners was carried out through a simple count of the number of organisations holding campaign events, which was reconstructed through examining internal reporting and media reports. The main challenge in measuring participation was that no comparisons were available, so it was difficult to assess the success of this aspect. For the OHCHR campaign in particular, the total number of potential partners, i.e. human rights NGOs, schools, government institutions, was unknown. For the ICRC campaign, the main partners were Red Cross and Red Crescent National Societies and therefore their total number was known and a

rate of participation could be calculated, even if no comparison was possible with previous campaigns of this nature.

The ICRC campaign also had an objective to increase organisational capacity to use new media and social networking (ICRC 3). Capacity is notoriously hard to evaluate and little rigorous evaluation of this has been carried out (Connolly & York, 2002; Medina-Borja & Triantis, 2007). In this regard, evaluating capacity was limited to the output level by assessing the usage of new media and social networking tools within the campaign through interviews with campaign staff and by observing the usage of these tools in the campaign. Limiting measurement of this point to a simpler less significant output level was another example of a compromise that was acceptable for the evaluation.

Media coverage of the campaign messages was examined for both campaigns (ICRC 4, 5 & 6, OHCHR 3 & 4). Media coverage was not set as a specific output for either campaign, but more so incorporated as additional indicators for awareness and attitude outcomes. This was not to consider coverage as a substitute for awareness measurement but a recognition that media coverage can be indicative of the place of an issue on the public agenda over time (Bauer, 2000). For this purpose, a quantitative media content analysis, a sub-set of content analysis (Macnamara, 2005) was carried out with the media analysis staff of both organisations. The media analysis adopted a co-occurrence analysis approach, which considered keywords in their context (Bauer, 2000). A coding scheme, a selection of keywords associated with the campaigns was established and automated searches carried out on third-party databases of a pre-selection of international/regional media and web news sources in English over the main 12 months period of each campaign. The analysis produced data illustrating the media coverage for both campaigns within the selected sample. This was useful in illustrating the ebbs and flow of media coverage throughout the campaigns and allowing comparison to non-campaign years. For example, with the OHCHR campaign, the number of mentions in the media on UDHR more than doubled in the campaign year, 2008, compared to the previous year, 2007 (2007, n=2598; 2008, n=6475).

The limitation to the media analyses carried out was that they were based on only a single factor (coverage or mention) and did not provide any semantic analysis (i.e. positive, neutral or negative) or context of the coverage (Macnamara, 2005). Secondly, the media analyses were limited to English-speaking media of the third-party databases so they could not be fully representative of the campaigns media coverage given that they both communicated in multiple languages.

The implementation of the methods to evaluate the output level results was found to be manageable and compromises made acceptable. This conforms to previous research that found that the communication output level more manageable to evaluate, although recognising their significance is less than outcome or impact level results (Cutlip, Centre & Broom, 2006; Lindenmann, 2003). However, the challenges seen in establishing a relatively simple measure, the levels of participation in comparison to an ideal level of participation proved difficult, particularly with the OHCHR campaign, reflecting the nature of global campaigning model used where ambitions may include using relays and reaching audiences that are not able to be fully specified at a campaign's creation (Manheim, 2011). Data collection on outputs was also reliant on the collaboration of the two organisations and their capacity to collect such data. For example, both organisations were asked to collate levels of campaign participation by their campaigns that was difficult to do comprehensively given the broad range of activities carried out (G. O'Neil, field notes, 4 December 2008 & 19 January 2009).

Compared to outputs, evaluating outcomes proved more challenging at different levels. In measuring changes to awareness and attitudes of targeted audiences in campaign evaluation, evaluators have traditionally relied on methods that canvassed directly target audiences. From a positivist perspective, this could be through pre-post survey (with a control group if feasible) of a sample of the target population (Valente, 2001). From a constructivist perspective, this could be through a participatory approach involving members of the target population defining the most significant changes from their perspective through methods such as outcome mapping (Carden, Smutylo & Earl, 2001). However, a main challenge with the global campaigning model used by the organisations

was the dispersed and diverse audiences being targeted which made canvassing of these audiences challenging. Alternative methods were therefore used, with varying levels of success, as described in the next paragraphs.

With the OHCHR campaign, a post-campaign survey of audiences was set up with a series of questions measuring awareness on issues addressed by the campaign (OHCHR 3 & 4). An online survey, established in French, English and Spanish was launched targeting a selection of countries whose national campaign partners (mainly NGOs and national human rights institutions) were encouraged by OHCHR to distribute it amongst their publics. However, the number of responses collected was insufficient, non-representative and not included in the evaluation findings. Limitations were seen in the ability of partners to collect data and for the evaluation to support and manage partners remotely. This was identified as a potential risk in the Methodology component as described above and was a reflection of the campaign model that relied on networks of partners as relays for communication (G. O'Neil, field notes, 4 December 2008).

For the ICRC campaign, a longitudinal panel study of young adults was set up. The panel was surveyed before, during and after the campaign. The panel, recruited from international university students studying in Switzerland, was asked firstly questions about themselves, their media habits, awareness and attitudes on humanitarian issues and then in the second and third surveys about the campaign itself (ICRC 4, 5 & 6). Although the panel showed progress in awareness over time (e.g. 31% awareness on the campaign in the third and final survey), these were not statistically significant differences (p values between 0.412 to 0.734). The likely explanation was a problem of attrition experienced with the panel, i.e. people dropping out from the 1st to 3rd panel survey (1st n=41; 2nd n= 36; 3rd n=26), which is a recognised methodological issue for panel studies (Elliot, Holland & Thomas, 2008). In addition to the panel, a one-off polling or "street survey" was undertaken of random residents during the campaign in Geneva, Switzerland, the birth place of the Red Cross where activities were being carried out, such as events, outdoor advertising and media placements. After exclusion criteria screening, 85

residents answered a short survey of 10 questions on their awareness of the campaign visuals and slogan showing awareness of 35%: 15% unprompted and 20% prompted.

The use of the surveying and polling posed several challenges. Firstly, both the panel and survey canvassed select target audiences and therefore could not claim to represent the views of all possible target publics of the campaign. Secondly, the ICRC panel was not able to find suitable replacements for members that dropped out leading to the issues of insignificant statistical results. The ability to use these methods was influenced by several factors. The resources available meant it was not possible to extend these research methods to more campaign audiences. Access to the audiences also posed issues, for example in being able to recruit further members of an appropriate profile for the panel. In discussions with the campaign managers, these limitations were recognised and were taken into account in how the data was then used in the findings presented (G. O'Neil, field notes, 25 May 2009).

The third group of effects focused on individual behaviour change (ICRC 7, 8, 9 & 10; OHCHR 5). Measuring behaviour change for communication campaigns usually relies on self-reporting of participants, observation or actual counts of behaviour (Grunig and Hunt, 1984), with the usual caveats around measuring behaviour in terms of self-assessment, reliability and intent (Dozier and Ehling, 1992). For the ICRC campaign, the focus of individual behaviour change was mainly on actions taken online to support the campaign (ICRC 9). This is a relatively new phenomenon in campaigning which has been labelled "clicktivism" (Karpf, 2010), that is, the act to support a cause or issue through the click of a computer mouse whilst on the Internet. The evaluator collaborated with the campaign manager and her staff on this data collection as statistics of website activities were collated by them from the campaign and third-party websites. The compilation of these statistics showed that some 150,000 persons undertook an action online, directly on the campaign website or on third party websites, such as the Cable News Network (CNN) website and on the social media network, Facebook.

Measuring online actions proved to be relatively straightforward; challenges were seen in the interpretation of the data as described below. The ongoing monitoring of these online actions also allowed the campaign to adjust its tactics and approaches accordingly (G. O'Neil, field notes, 25 May 2009). However, when it came to data collection, several objectives and their indicators on individual behaviour change proved to be too impractical for the evaluation to measure, for example the ICRC campaign to "motivated people to undertake voluntary work" (ICRC 8). In discussions with the organisation, no procedures were found to implement methods that could measure this outcome given its broad goal. Alternative or proxy measures were therefore considered that provided useful but incomplete data. For example, in the rolling survey of National Societies by their International Federation, they asked if increased requests for volunteering had been noted during the campaign period (G. O'Neil, field notes, 25 May 2009).

The final group of effects focused on influencing change in institutions (ICRC 10, OHCHR 5 (partially), 6 & 7). Both campaigns sought to influence the raising of funds for their respective institutions (ICRC 10, OHCHR 7). The only available measurement of fundraising for the ICRC campaign was through statistics of the online donation action which showed no substantial increase in donations for the campaign year (G. O'Neil, field notes, 28 January 2010). The OHCHR campaign focused on attracting additional funds from governments and set up a financial appeal for this purpose. According to OHCHR monitoring records, this appeal raised 1.1 million US dollars and attracted nine governments to donate for the first time in 2008 (OHCHR, 2008). Output/outcome 5 of the OHCHR campaign included "reaffirmed commitment of governments, civil society, educational, cultural and human rights institutions to values and principles of UDHR". As no specific tracking system was in place, the evaluator worked with the campaign manager and his staff using media reports and internal reports to collate and categorise retroactively any publicised commitments made by governmental institutions and human rights activists. For output/outcome 6, "Helped bridge gaps in HR implementation at the national level", the main indicator used to assess progress was the ratification or signature by countries of the 25 international treaties centred on human rights. This was measured

through the public records kept by the treaty bodies on the ratifications and signatures which found no major increase in the campaign year (2008).

Data collection on institution change was influenced by several factors. The network approach to campaigning as adopted by both organisations, limited the ability to collect all relevant data from partners that were acting as relays for the campaigns. For example, access to data on funds raised by the main campaign partners for the ICRC, the 186 National Red Cross and Red Crescent Societies was not accessible. Although members of the same movement, the National Societies are independent organisations, making a centralised monitoring system for fundraising not feasible. For both organisations, it emerged during the data collection that the monitoring systems for such indicators were under-developed. As a result, the evaluator worked with the campaign managers and their staff to develop some rudimentary tracking systems that proved only partially sufficient in measuring progress on the indicators. (G. O'Neil, field notes, 13 March 2009 & 19 May 2009).

Before proceeding to the data analysis stage, the evaluator sought to involve campaign staff and partners of both organisations. This was carried out through both semi-structured interviews with campaign staff and partners (11 persons for OHCHR and 25 for the ICRC) and surveys with campaign partners (247 respondents for the OHCHR and 52 for the ICRC). These interviews and surveys were carried out with the purpose of both involving staff and partners in the evaluation and triangulating their feedback with the findings from the other data sources, an approach also to strengthen the reliability of the data collected (Lincoln & Guba, 1985). This was seen as a complement but not a substitute for data collected directly from audiences. However, given the challenges of access to audiences, these methods supported the evaluators in understanding where staff and partners considered progress was made and what challenges were seen. For example, with the ICRC campaign, partners reported that they struggled to deal with the six competing messages of the campaign and the wide range of activities they could participate in. For the OHCHR campaign, partners reported seeing an added value in

being associated with a global campaign although they lacked guidance on all messages and activities in order to be fully active.

Further, in discussions with staff, their perception and expectations of the evaluation emerged which influenced the evaluator in the analysis of the data collected. For example, the main focus of the evaluations was on evaluating the outputs/outcomes of the campaigns; whereas the interest of the staff of both organisations was in learning how to improve the efficiency of their activities in contributing to the outputs/outcomes. This perspective was then taken into account and the eventual evaluation findings were re-shaped accordingly.   Involving staff and partners were also part of the collaborative approach to develop shared ownership, build confidence in the evaluation findings and encourage eventual learning from the evaluation process (G. O'Neil, field notes, 4 December 2008 & 19 January 2009).

Data was analysed as the second step of the Implementation component. For both campaigns, data collected for the different indicators (17 indicators for OHCHR and 18 for ICRC) was analysed using appropriate analytical methods including regression modelling of quantitative data, qualitative analysis of interview transcripts, coding and analysis of media coverage and the collation of various monitoring data. Given the number of indicators and research methods used, the data generated was considerable. As a response, the evaluator referred to the theory-based model of the evaluations. The data was structured and organised to understand the extent to which the campaign activities had produced outputs that then contributed to outcomes and the overall objectives, an essential analytical strategy of this model (Fitz-Gibbon & Morris, 1996). This was summarised in the draft evaluation reports in the form of an assessment of each of the campaign outputs/outcomes that then led to subsequent conclusions and recommendations of the evaluator (G. O'Neil, record sheet, 7 July 2009 & 2 June 2010).

In analysing the data and presenting the findings, the evaluator was careful to present the evidence to justify the findings made and any claims of casual inference, that is, the demonstration that a campaign led to a given change. In some instances, this was not an

issue, such as in publics taking online action for the ICRC campaign or OHCHR partners becoming involved in the campaign, as in both cases the changes could be shown to have been directly triggered by the campaigns. In other instances, it proved harder to establish such claims, such as the commitments made to human rights for OHCHR campaign or the change in attitudes of publics for the ICRC campaign where data was not collected or the other possible influences on commitments or attitudes were potentially many. These limitations were discussed with the organisations and reflected in the findings of the two evaluation reports where the strength of any causal claims was explained (G. O'Neil, record sheet, 7 July 2009 & 2 June 2010).

Two examples are provided to illustrate the challenges seen in data analysis. Issues were seen in understanding and interpreting the significance of what had been measured for some data. For example, for the measuring of online actions for the ICRC campaign, that numbered some 150,000, the conversion of this relatively superficial online actions ("a click") into a substantial action as desired by the campaign ("a humanitarian gesture") was low (reported by 537 participants on the campaign website) although a limitation could be that actions were unreported. The challenge to convert "clicktivism" into more concrete action has been observed previously in online campaigns (Morozov, 2009) where it has been argued that online actions do not lead to more in-depth "offline" action but more so are part of shaping individuals' online identities (Zhao, Grasmuck & Martin, 2008). This was also discussed and debated with the campaign manager and her staff, with different views as to the significance of such online actions. At the time of the evaluation, the communication evaluation field provided limited guidance on these issues (G. O'Neil, field notes, 7 July 2010).

Issues were seen in the complexity and time taken to see changes as a result of the campaigns. For example, with the OHCHR campaign, the lack of progress seen in ratification or signatures of treaties illustrated the limits of measuring the results in only the campaign period (one year) or directly after, given that the trajectory of change for such influence may not be immediate or direct (Woolcock, 2009). Studies show that a government's decision to ratify a treaty can be proceeded by other actions and the actual

act of signature or ratification may not be seen immediately (Goodman & Jinks, 2003). In this regard, it has been suggested that indicators earlier in the implementation process would be more appropriate (Starling, 2010) and a recognition that long term effects, such as treaty ratification, may not be visible in the campaign period or directly after (Salmon, 1989).

During the Implementation component, the role of the evaluator varied in the different steps, as proposed by Rossi et al. (2007). In data collection, the evaluator was a "facilitator", assessing data as it became available and involved in an iterative process continuing a dialogue with the staff of both organisations on the results seen and integrating their feedback and reactions. In the data analysis, the author played more the role of "judge" in assessing the findings and drawing conclusions independently. When these findings were discussed with the organisations, the author resumed the role of "facilitator" in discussing and refining conclusions and recommendations based on the feedback of campaign staff and any new input they provided.

Overall, the analysis and reflections of the Implementation component builds on the scenario described in the Methodology component: the methodological choices were implemented in some cases playing out as expected and others not. The deployment of data collection tools for the evaluation faced issues that were mostly linked to the global campaign model: the diversity of objectives; the difficult to identify and access audiences; the network nature of partners and the challenges to work with them in data collection.

The implementation component showed how the evaluator needed to adapt certain aspects, abandoning some methods and indicators, seek plausible and acceptable alternatives and understand the significance of what was being evaluated; a mixture of elements resembling the "muddled middle" (p. 177) as envisaged by Patton (2011). This situation resembles also the description given by Tourmen (2009) in her study of actual evaluation practice in France:

> This activity consists of major and minor choices, back-and-forth movements, difficulties in decision making, compromises, contexts that are not completely under control nor easy to foresee, and so on. (p.28)

The analysis of data, as carried out solely by the evaluator, was the most isolated element of the evaluation process but still validation and contribution were sought from staff, reflecting the participative approach adopted. On this basis, the evaluator was able to provide an assessment on the given outputs/outcomes whose evidence needed to be transparent in explaining the limitations seen and allowing the organisations to judge for themselves the strength of the findings and their basis for the consequent conclusions and recommendations.

5.6.3. Similarities and differences between the ICRC and OHCHR

The analysis and reflection on Evaluation and Implementation components highlights both similarities and differences between the organisations, their campaigns and the consequent evaluation carried out. The communication function of the ICRC was larger than that of the OHCHR in terms of staff and resources, which meant that more consultation and discussions were needed with ICRC staff in the evaluation process, as illustrated in Figure III. The similarities of the campaigns in their global ambitions and model meant that similar methodological approach and procedures could be adopted. The largest difference seen was in the implications of the choice of the evaluation design between pre-post (ICRC) and post-only (OHCHR) as described above. The experience of the evaluator was also different in respect of the two organisations; with more familiarity and experience with the ICRC than with OHCHR.

5.7. Discussion: choices, value and influence

The experience of evaluating the two campaigns provided an analysis and reflection on the methodology and implementation components of the evaluation process. The insights this provided as to what is appropriate and feasible for evaluation methodology for

communication activities of IOs is now discussed followed by consideration of the internal and external factors that influenced implementation.

5.7.1. Methodological choices

The methodological choices made and the procedures put in place, rather than being shaped by a dominant evaluation paradigm of the organisations or the evaluator, were more so constructed through a joint view considering the campaign model, organisational settings and contexts. In doing so, the methodological approach resembled a relatively recent view that has been labeled the "Eclectic-Mixed Methods-Pragmatic Paradigm" (Reeves, 1997); "eclectic" referring to the combination of different methods and "pragmatic" in recognizing the need to adapt and change methods and procedures in their implementation. This approach also seeks to find an accommodation between the linear positivist structure of the evaluation framework and the complexity it encounters, which in this case, was the campaign models and the contexts in which they were used. This perspective emerged during the evaluations and was shaped by the interaction between the Methodology and Implementation components, illustrating both its feasibility and appropriateness.

Derived from this perspective, the choice of the evaluation design proved to be key. The use of the pre-post evaluation design for the ICRC campaign strengthened the evaluation, such as the ability to use more robust methods and the possibility to influence campaign design and implementation. The post-only design used for the OHCHR campaign presented certain disadvantages described above which correspond with previous findings on post-only designs (Broom & Dozier, 1990; Lehmann, 2007). However, the possibility to use a pre-post evaluation design was reliant on evaluation being considered and integrated within the communication function early in the campaign planning process. The use of experimental and quasi-experimental designs (pre-post with control groups) was not found to be feasible for the model of campaigning used in these two organisations. Exposure was desired across multiple audiences with different purposes, making the use of such designs impractical, as already found previously (Broom & Dozier, 1990; Flay & Cook, 1981).

The nested analysis mixed methods strategy, which shaped the evaluation framework and the selection of methods, provided an appropriate way of organising the evaluation in manageable sub-sets of outputs/outcomes of which a range of methods and indicators could then be matched. This methodology effectively broke down the campaign objectives into smaller elements that could then be "nested" in an overall framework. Challenges were seen in the ability to match all indicators with appropriate methods that could then be implemented. This did not put into question the nested strategy, but more so the ability to evaluate the full range of outputs/outcomes of the campaigns.

At the same time, the theory-based model guided the strategy and the consequent data analysis. The use of the theory of change proved to be appropriate but the experience showed that the theories were constructed at a high level and did not explore sufficiently the potential pathways to change and how to capture this. For example, in understanding the different steps to treaty ratification of the OHCHR campaign or the steps of online action to more significant change in the ICRC campaign. Consulting previous research on campaign models (Hwang, 2012), theories of behaviour change (Synder, 2007) and communication effects (Macnamara, 2006) could have made for more robust theories of change that underpinned the campaigns and modified expectations about what could be achieved and how. The absence of theoretical aspects being considered in the design of communication campaigns has been previously reported (Dozier, Grunig & Grunig, 2001; Macnamara, 2006).

Methods for output were largely able to evaluate this level, although issues were seen with the ability of the organisations and the evaluator to monitor all partner activities. In addition, the reliance on partners for data collection proved difficult, also at the outcome level, for example in surveying audiences through partners. At the outcome level, methods to canvas target audiences directly such as surveys, polls and panel studies were found only to be feasible for accessible, limited and defined audiences.

Tracking various data to measure individual and institutional change proved feasible; challenges were more so seen in the definition of the appropriate point in a given pathway of change to measure (e.g. tracking treaty ratification) or the interpretation of what was being measured (e.g. tracking online actions). The ability to understand and analyse the information collected was supported by the participative approach adopted that involved interviewing and surveying campaign staff and partners. This strengthened the data collected and also developed the confidence of campaign staff in the evaluation and its findings.

To evaluate more in-depth the campaign outcomes and strengthen the findings, the evaluations could have used additional methods. For example, the literature cites multi-country omnibus surveys and propensity score matching of survey data for measuring changes to knowledge and attitudes, case studies and outcome mapping to address more unknown or non-linear outcomes, contribution analysis or process tracing to consider the influence on policy and institutions and the emerging field of web analytics to observe and monitor online behaviour (Beach & Pedersen, 2013; Gonçalves & Ramasco, 2008; Lindenmann, 2003, O'Neil, 2013; White, 2005). These methods would fit within a nested analysis mixed methods strategy and would be appropriate for global campaigns. However, it was more a question of feasibility for this study in that both organisations sought a broad view of all outputs/outcomes and the use of such methods would require an emphasis on some outputs/outcomes over others, given resource limitations.

Ultimately, the methodological choices made and their implementation allowed the evaluator to break down the campaigns into sub-sets of outputs/outcomes for which findings were produced, with conclusions reached per output/outcome rather than globally for the campaigns (G. O'Neil, record sheet, 7 July 2009 & 2 June 2010). The collaborative approach with campaign managers and their staff allowed them to assess the strength of the data and the claims made, as they participated in decisions taken on methodological choices, saw the challenges seen in their implementation and the limitations. This involvement of managers and staff was seen as key so that they saw that

the level of methodological rigor achieved was a result of choices made jointly (Braverman & Arnold, 2008).

5.7.2. Value of evaluation

Moving from the Implementation to the Findings and Use components, it has been put forward that the level of use of an evaluation is a strong indication of its value (Patton, 2008). Based on the initial reception of the evaluation reports in both organisations as witnessed by the evaluator in the findings workshops, the findings, conclusions and recommendations of the evaluation reports were considered seriously by the campaign managers and their staff, provoking discussions and reflections, notably on what could be learnt for future communication campaigns and programmes (G. O'Neil, field notes, 8 October 2009 and 17 September 2010).

To understand further the value of the evaluations from the perspective of their actual use would require a more in-depth study, which is the focus of the third article of this PhD. However, what this author could observe while interacting with the staff of organisations was how their interest in the evaluation developed and the opportunity for learning through the evaluation process was present (G. O'Neil, field notes, 4, 12 & 26 March 2009; 7, 19 & 25 May 2009, 26 August 2009, 15 & 28 January 2010). This possibility, that involvement in the evaluation process provokes learning and other changes, has been previously recognised in the literature and is labelled as "process use", (Alkin & Taut, 2002; Patton; 2000).

5.7.3 Influence of internal and external factors

Reviewing the methodology adopted and implemented for the two campaign evaluations, the preceding paragraphs provided insights into what were both suitable and in practice possible to do for both organisations. The picture painted is a pragmatic and adaptive evaluation approach that provided broad findings on all campaign outputs/outcomes but compromised the range of methods that could be used and the consequent strength of the findings. This was balanced by the participative nature of the evaluation that built

confidence in the evaluation findings and laid the ground as to how they would be received and used by the campaign staff.

The discussion also touched on factors outside of the evaluation process that influenced the methodology adopted and implemented for the two campaign evaluations. These factors are now considered further, with reference to the evaluation framework described above (Figure I).

The goals and ambitions of the communication campaigns clearly shaped the evaluation methodology and its implementation. As described above, the consideration and integration of evaluation in the campaign planning process directly influenced the choice of the evaluation design, and thus its robustness, which corresponds to the findings of the broader study of communication evaluation of IOs carried out by this author (O'Neil, 2013). The global campaigning model adopted by both organisations, with multiple objectives targeting various dispersed audiences with different effects sought impacted on the ability of the evaluation to match this with appropriate indicators, methods and means to collect the relevant data. Dozier (1990) contended it was the complexity of evaluation methodology that made it difficult to apply to communication activities. However, through the experience of evaluating these two campaigns this author found the contrary: it was more so the complexity of the campaign models that proved challenging to apply the appropriate methodology and methods.

Communication goals and ambitions were also found to be a reflection of the organisational settings and contexts; in that the campaigns needed to reflect the priorities of the organisations. Positively, this could influence their relevance to the organisations but at the same time it meant addressing diverse and dispersed audiences that impacted on the ability to evaluate as described above (G. O'Neil, field notes, 29 May 2008 & 25 June 2008).

The structure of both organisations meant that both campaign managers and their staff were integrated within the communication function that was part of an overall

management structure. For example, both campaigns had to have their global campaign strategies and budgets approved by senior management committees (G. O'Neil, field notes, 29 May 2008 & 25 June 2008). However, once strategy was approved, the organisational settings provided the campaign managers with some flexibility to adapt activities and objectives in reaction to the opportunities that arose and the changing contexts. In the case of the ICRC, given the pre-post evaluation design, this meant that there was a possibility for the campaign team to adapt and react to the evaluation findings as they emerged, and did so in several instances, such as issues of messaging or take-up of activities, but this was also due to the monitoring data they were collecting themselves, such as media coverage reports and website statistics. For both organisations, this flexibility also meant they could initiate and manage evaluations themselves.

The nature of the organisation's global structure also influenced the campaigns' implementation and the consequent evaluation approach adopted. Both organisations had head offices where their communication functions and campaign staff were located and a series of field offices spread around the world, combined with a network of partners, that both served as campaign relays, as described in Table II. Therefore, the organisational structure used for campaigning was less of a hierarchical form but more of a combined confederation and network form that has been shown to be highly effective in global campaigning (Keck and Sikkink, 1998; Manheim, 2011). However, while effective for campaigning, the network structure brought with it particular challenges for the evaluations, such as the difficulty of the evaluation to support partners in carrying out evaluation tasks, having access to monitoring data they collected and direct access to their audiences, as described above. This also reflects challenges seen previously in international-level evaluation, distinguishing it from national-level evaluation (Mathiason, 2011). Issues of access to data and audiences were also indicated in the broader study of communication evaluation of IOs carried out by this author (O'Neil, 2013).

There was limited input from the central evaluation policy and institution into the campaign methodologies and its implementation given that the evaluations were commissioned directly by the communication functions and not by the central evaluation

units. Centralised evaluation policies of both organisations were also in development at the time of the evaluations (G. O'Neil, field notes, 29 May 2008 & 25 June 2008). Therefore, it was found that evaluation had not yet reached a stage of maturity or institutionalisation as seen in other organisations (Dahler-Larsen, 2012; Schwandt, 2009), which in the case of the EU has found to be a positive influence in making evaluation more systematic and routine (Højlund, 2014b). More so, in the case of these two evaluations, the main influence of the evaluation institutions and their policies were indirect. This was seen in the terminology, procedures and framework of the evaluations adopted and approved by the campaign managers that largely mirrored the institutional approach as reflected in the RBM systems in place, which was the most visible element of evaluation policies and institutions for staff.

The availability of the necessary budgets has consistently been found to be a barrier to communication evaluation across all sectors (Cutlip et al. 1994; Macnamara, 2006; Watson & Noble, 2007). Within these two campaigns, the evaluator worked with available resources and was conscious that budgets were not available for methods that could have strengthened the evaluation, such as launching multi-country omnibus surveys or multi-site visits for case studies.

The question of know-how and absence of an evaluation culture amongst communication staff has been raised as an obstacle for communication evaluation (Watson, 1995; Macnamara, 2006). Within the two organisations, campaign staff were familiar with evaluation concepts such as setting measureable objectives and indicators, due to the RBM systems and the emerging evaluation institution and policies, as mentioned above. However, while the communications unit of the ICRC had carried out smaller-scale evaluations previously, it was the first documented evaluation of a global campaign for both organisations (G. O'Neil, field notes, 29 May 2008 & 25 June 2008).

What was confirmed by these evaluations was the reliance on external persons, in this case, this author (the evaluator), to carry out the bulk of evaluation tasks. This was not due to the lack of qualified staff but more so to the limited integration of evaluation into

the schedules and tasks of staff (G. O'Neil, field notes, 29 May 2008 & 25 June 2008). This had the advantage of producing independent external evaluation findings, which has traditionally been seen as a key attribute of evaluation (Pawson & Tilley, 1997; Weiss, 1998) but had the disadvantage of a low integration of the evaluation element within the communication function, with implications for the design and methods adopted – and possibly detrimental as to how the results would be considered and used by the relevant communication staff (Mendelsohn, 1973). This reliance on external persons and low integration of evaluation within the communication function was also confirmed in the wider study of IOs by this author (O'Neil, 2013). Coping strategies were developed and implemented to counter this low integration, mainly the participative nature of the evaluations as described above.

The context in which IOs carry out their campaigns is a global setting with broad and diverse audiences, which implies a level of complexity and unpredictability of response to communication activities (Lennie & Tacchi, 2013; Manheim, 2011). The two campaigns adopted various strategies to cope with these contextual issues which had consequences for the evaluations. For example, mass targeting of audiences was carried out in order to produce active segments of the audience, a strategy previously documented (Dozier et al., 2001). For the evaluations, challenges were seen in its ability to canvas large potential audiences and use appropriate methods to reach those segments that became active. However, context in many respects was an indirect influence on the evaluations. Context directly influenced the organisational settings and how they responded through actions such as the communication campaigns, which in turn influenced the campaign model adopted and the evaluation approach used, a representation of the interrelations within organisations and externally, as recognised in systems theory (Dubin, 1976; Williams & Imam, 2007).

The evaluation and communication evaluation fields were relevant external factors. The methodology and its implementation as developed by the evaluation in collaboration with the campaign managers drew from these fields to various extents, implicit in most cases. For example, the linear logical evaluation pathway used (from activities to outcomes) has

been widely promulgated by both fields (CCMC, 2004; Lindenmann, 2003; Shaw, Greene & Mark, 2006; Weiss, 1998). At the same time, the evaluation institutions and policies that had some influence, such as the RBM systems, were also largely consistent with the influence of these fields (G. O'Neil, field notes, 4 March 2009 & 7 May 2009). Aside from these broad conceptual notions, there was no specific influence from the relevant fields, for example in an expectation of the campaign managers to use an experimental and quasi-experimental design as seen in most campaign evaluations to date (G. O'Neil, field notes, 29 May 2008 & 25 June 2008).   More so, the mixed methods approach adopted was more reflective of the practice promoted by the communication evaluation field (Broom & Dozier, 1990; Michaelson & Macleod, 2007).

It was found that factors could operate both as "enabling" or "impeding" factors (Højlund, 2014b; Mark & Henry, 2004). For example, the organisational setting impeded the evaluation through the network/confederation structure used which brought with it particular challenges as described above. At the same time, the organisational setting provided flexibility to the campaign staff in commissioning and managing the evaluation that then facilitated its implementation. This dual enabling/impeding role was also seen for communication goals and ambitions.

In summary, internal factors rather than external factors were found to be more important in influencing the methodology adopted and implemented for the two campaign evaluations. The scope of communication goals and ambitions, the campaign model, the level of integration of evaluation within the communication function, the organisational setting and structure adopted were highlighted. Arguably, some of these factors were shaped in reaction to the global contexts within which these organisations operated.

5.7.4. Limitations

This article itself expands considerably upon the limitations faced in evaluating the two campaigns. Further, the two organisations of this research were not selected randomly from the international public sector, more so they were both willing to collaborate on this research. As a consequence, their representativeness of the whole sector is limited. As

this article is a reflection on the research carried out by this author from the "insider-outsider" perspective, there are potential issues of bias in the analysis that have previously been seen with this type of research (Dwyer & Buckle, 2009; Kerstetter, 2012).

5.8. Conclusions: towards pragmatic and adaptive evaluation

This article sought to understand the challenges in implementing communication evaluation methodologies by providing an analysis and reflection on the carrying out of the evaluation of two communication campaigns.

Overall, a pragmatic and adaptive evaluation approach emerged as an appropriate response given the ambitions of the campaigns and the challenges raised by the factors of influence identified. This approach provided broad findings on all campaign outputs/outcomes and learnings on improving the efficiency of campaign activities, involving campaign staff in the evaluation process which contributed to their confidence in the evaluation methods and eventual findings.

Evaluations of this nature could be further strengthened through ensuring the use of an appropriate evaluation design (pre-post) and further diversity in the methods used. However, this is largely dependent upon addressing the challenges posed by factors outside of the evaluation process, such as the scope of communication goals and ambitions, the campaign model used, the level of integration of evaluation within the communication function, the organisational settings and context. Therefore, although organisations could adopt appropriate evaluation methodologies, designs and methods they also have to consider the feasibility of countering and/or capitalising on these outside factors in order to facilitate communication evaluation that is both robust and of value to organisations.

The indications are that organisations will increase their use of communication for strategic purposes and reinforce their evaluation of campaigns and other communication activities (Likely & Watson, 2013; Michaelson & Macleod, 2007; Schwarz & Fritsch, 2014). Further, it can be predicted that the IO campaign model of multiple outcomes,

activities, partners and audiences will remain dominant, given that campaigns need to support the ambitions of these organisations to maintain their influence and place in the world. At the same time, evaluation will most likely have the same resources available and given the expectations seen in these two cases, it will be of more value when focused on providing broad findings to improve the efficiency of campaign activities rather than in-depth examination of particular outputs or outcomes. However, value needs to be assessed beyond the methodological choices and their implementation; best placed it seems would be by those who commission and use evaluation findings as suggested in the literature (Bamberger et al., 2011; Braverman & Arnold, 2008; Patton, 2008).

One caveat should be noted. Organisations are moving rapidly to place greater emphasis on the online environment for their communication activities (Zerfass et al., 2012) and this area has more potential for evaluation, possibly altering the challenges seen to date. As a consequence, evaluation will need to adapt by proposing appropriate strategies and methods for this new emerging area. The experience of evaluating the online actions within the ICRC campaign illustrated the challenges faced in interpreting and understanding the significance of this environment and in the given context.

The evaluations undertaken were for two IGOs rather than INGOs. Arguably, the conclusions of this article also apply to INGOs as campaigning is similar within both types of organisations with one distinct difference: As IGOs are inter-governmental institutions, i.e. created by governments they are possibly less able to adopt goals that would potentially put them in conflict with governments, whereas INGOs are in theory able to undertake such campaigning. Then again, this distinction is increasingly blurred as it has been remarked that IGOs now have their own agendas independent of the governments that created them (Barnett & Finnemore, 1999).

A number of findings of this article have implications beyond the international public sector. The experience of carrying out communication evaluation and the challenges faced in implementation has had limited attention, with the findings of this article relevant for the broader communication evaluation and evaluation fields. Further, this

article is a contribution to the body of "insider-outsider" research which has seen limited studies in organisational settings such as IOs.

Finally, this article considers the campaign level of activities which can be extended to a broader range of activities, for example at the programme level. However, given the ambitions of these organisations, and their altruistic goals as reflected in their campaign objectives, further research would be required on their contribution beyond the campaign and programme level, that is, at the organisational and societal levels.

5.9. References

Alkin, M. C., & Taut, S. M. (2002). Unbundling evaluation use. *Studies in Educational Evaluation*, 29(1), 1-12.

Alvarez, J. (2006). International Organisations: Then and Now. *The American Journal of International Law, 100*, (2), 324-347.

Ayers, T. D. (1987). Stakeholders as partners in evaluation: A stakeholder-collaborative approach. *Evaluation and Program Planning*, 10, 263-271.

Bamberger, M., Rugh, J., & Mabry, L. (2011). *RealWorld evaluation: Working under budget, time, data, and political constraints*. Newbury Park, CA: Sage Publications.

Barnett, M., & Finnemore, M. (1999). The Politics, Power and Pathologies of International Organizations. *International Organizations.* Vol. 53, No. 4: 699-732.

Bauer, M. W. (2000). Classical content analysis: a review. In M. W. Bauer & G. Gaskell (Eds.), *Qualitative analysis with text, image and sound* (pp. 131-251). London: Sage Publications.

Beach, D., & Pedersen, R. B. (2013). *Process-Tracing Methods: Foundations and Guidelines*. Ann Arbor: University of Michigan Press.

Braverman, M. T., & Arnold, M. E. (2008). An evaluator's balancing act: Making decisions about methodological rigor. In M. T. Braverman, M. Engle, M. E. Arnold & R. A. Rennekamp (Eds.), *Program evaluation in a complex organizational system: Lessons from Cooperative Extension*. *New Directions for Evaluation*, 120, 71–86.

Broom, G., & Dozier, D. (1990). *Using research in public relations*. Upper Saddle River, NJ: Prentice Hall.

Campbell, D. T. (1984). Can we be scientific in applied social science. *Evaluation studies review annual*, Volume 9. Beverly Hills, CA: Sage Publications.

Carden, F., Smutylo, T., & Earl, S. (2001). *Outcome mapping: Building learning and reflection into development programs*. International Development Research Centre.

Charnovitz, S. (2006). Nongovernmental Organisations and International Law. *The American Journal of International Law*, *100*(2), 348-372.

Chesterton, P. (2004). *Evaluation of the Meena Communication Initiative*, UNICEF Regional Office for South Asia, Kathmandu.

Christie, C., & Fierro, L. (2012). Evaluation policy to implementation: An evaluation of scientifically based research in practice. *Studies in Educational Evaluation*, 38: 65–72.

Coffman, J. (2002). *Public communication campaign evaluation: An environmental scan of challenges, criticisms, practice, and opportunities*. Cambridge, MA: Harvard

Coffman, J. (2003). *Lessons in evaluating communications campaigns: Five case studies*. Harvard Family Research Project. Retrieved from:
http://www.hfrp.org/content/download/1114/48619/file/lessons_comm_campaigns.pdf

Communications Consortium Media Center. (2004). *Guidelines for Evaluating Nonprofit Communications Efforts*. Retrieved from:

http://www.mediaevaluationproject.org/paper5.pdf

Connolly, P., & York, P. (2002). Evaluating capacity-building efforts for nonprofit organizations. *OD Practitioner*, *34*(4), 33-39.

Cousins, J. B. (1995). Assessing program needs using participatory evaluation: A comparison of high and marginal success cases. In J.B. Cousins & L. M. Earl (Eds.), *Participatory evaluation in education: Studies in evaluation use and organizational learning* (pp. 55-71). London: Falmer.

Cousins, J., & Chouinard, J. (2012). *Participatory evaluation up close: A review and integration of research-based knowledge*. Charlotte, NC: Information Age Press.

Cutler, A. (2004). Methodical failure: The use of case study method by public relations researchers. *Public Relations Review*, 30(3), 365-375.

Cutlip, S. M., Center, A. H., & Broom, G. M. (2006). *Effective public relations* (9th ed.). Upper Saddle River, NJ: Prentice Hall

Dahler-Larsen, P. (2012). *The Evaluation Society*. Paolo Alta, CA: Stanford Press.

Datta, L. (1994). Paradigm wars: A basis for peaceful coexistence and beyond. *New Directions for Program Evaluatio*n, 61: 53-70.

Dozier, D. (1990). The innovation of research in public relations practice: Review of a program of studies. *Public Relations Research Annual, 2*, 3-28.

Dozier, D., & Ehling, W. (1992) Evaluation of PR Programs: What the Literature Tells Us About Their Effects. In J. Grunig (Ed.), *Excellence in Public Relations &*

Communication Management (pp. 159-184). Hillsdale, NJ: Lawrence Erlbaum Associates.

Dozier, D., Grunig, L., & Grunig, J. (2001). Public Relations as Communications Campaign. In R. Rice & C. Atkins (Eds.), *Public communication campaigns 3rd Edition* (pp. 231-248). Newbury Park, CA: Sage Publications.

Dubey, D. & Bardhan, A. (1981). *Evaluation of communication programmes*. UNESCO Communication Research Module No. 5. New Delhi: Caxton Press.

Dubin, R. (1976). Theory building in applied areas. In M.D. Dunnette (Ed.), *Handbook of Industrial/organizational psychology* (pp. 17-39). New York, NY: Rand McNally.

Dwyer, S. C., & Buckle, J.L. (2009). The Space Between: On Being an Insider-Outsider in Qualitative Research. *International Journal of Qualitative Methods* 8:54–63.

Elliott, J., Holland, J., & Thomson, R. (2008). Longitudinal and Panel Studies. In P. Alasuutari, L. Bickman & J. Brannen, *The SAGE Handbook of Social Research Methods* (pp. 228–249). Thousand Oaks, CA: Sage Publications.

Elman, C., & Elman, M. (Eds.) (2003). *Progress in International Relations Theory*. Cambridge, MA: MIT Press.

Flay, B., & Cook, T. (1981). Evaluation of Mass Media Prevention Campaigns. In R. Rice & W. Paisley, *Public Communication Campaigns* (pp. 239–264). London: Sage Publications.

Fischer, R. (1995). Control construct design in evaluating campaigns. *Public Relations Review*, *21*(1), 45-58.

Fitz-Gibbon, C. T., & Morris, L. L. (1996). Theory-based evaluation. *Evaluation Practice*, 17(2), 177-184.

Fitzpatrick, J. L. (2012). An introduction to context and its role in evaluation practice. *New Directions for Evaluation*, 2012(135), 7-24.

Funnell, S., & Rogers, P. (2011) *Purposeful Programme Theory. Effective use of theories of change and logic models*. San Francisco: Jossey-Bass.

Gonçalves B., & Ramasco J. (2008). Human dynamics revealed through Web analytics. *Physical Review*, 78, 1-7.

Goodman, R., & Jinks, D. (2003). Measuring the Effects of Human Rights Treaties. *European Journal of International Law*, 14(1), 171-183.

Greene, J. G. (1988). Stakeholder participation and utilization in program evaluation. *Evaluation Review*, (12), 912-116.

Greene, J., Mark, M., & Shaw, I. (2006). *Handbook of Evaluation: Policies, Programs and Practices*. London: Sage Publications.

Gregory, A., Morgan, L., & Kelly, D. (2005). *Patterns of PR in Britain's 'most admired' companies and public sector organisations*. Leeds: Leeds Business School for Eloqui Public Relations.

Gregory, A., & White, J. (2008). Introducing the Chartered Institute of Public Relations' initiative. In B. van Ruler, A. Tkalac-Verčič & D. Verčič (Eds.), *Public relations metrics* (pp. 307–17). New York, NY: Routledge.

Grunig, J. (ed.) (1992). *Excellence in Public Relations and Communications Management.* Hillsdale, NJ: Lawrence Erlbaum Associates

Grunig, J. (2008). Conceptualizing quantitative research in public relations. In B. van Ruler, A. Tkalac-Verčič & D. Verčič (Eds.), *Public relations metrics* (pp. 88-119). New York, NY: Routledge.

Grunig, J. (2009). Paradigms of global public relations in an age of digitalisation. *PRism,* 6(2). Retrieved from:
http://www.prismjournal.org/fileadmin/Praxis/Files/globalPR/GRUNIG.pdf

Grunig, J., & Hunt, T. (1984). *Managing Public Relations*. Fort Worth, TX: HBJ Publishers.

Henry, G. T., & Gordon, C. S. (2003). Driving less for better air: Impacts of a public information campaign. *Journal of Policy Analysis and Management*, 22: 45–63.

Højlund, S. (2014a). Evaluation use in the organizational context – changing focus to improve theory. *Evaluation.* Vol. 20(1): 26–43.

Højlund, S. (2014b). Evaluation use in evaluation systems–the case of the European Commission. *Evaluation,* 20(4), 428-446.

Hon, L. (1998). Demonstrating Effectiveness in Public Relations: Goals, Objectives, and Evaluation. *Journal of Public Relations Research*, 10, 103–135.

Hwang, Y. (2012). Social Diffusion of Campaign Effects: Campaign-Generated Interpersonal Communication as a Mediator of Anti-tobacco Campaign Effects. *Communication Research*, 39 (1), 120-141.

Hyman, H. H., & Sheatsley, P. B. (1947). Some reasons why information campaigns fail. *Public Opinion Quarterly*, *11*(3), 412-423.

International Committee of the Red Cross. (2010). *About the International Committee of the Red Cross.* Retrieved from: http://www.icrc.org/eng/who-we-are/overview-who-we-are.htm

International Organisation for Migration. (2002). *Awareness Raising and Information Strategy on People in Need of International Protection - Evaluation Report*.

Jelen, A. (2008). The nature of scholarly endeavours in Public Relations. In B. van Ruler, A. Tkalac-Verčič & D. Verčič (Eds.), *Public relations metrics* (pp. 36-59). New York, NY: Routledge.

Kaldor, M., Moore, H., & Selchow, S. (Eds.). (2012). *Global Civil Society 2012: Ten Years of Critical Reflection.* Basingstoke: Palgrave Macmillan.

Karpf, D. (2010). Online Political Mobilization from the Advocacy Group's Perspective: Looking Beyond Clicktivism. *Policy & Internet,* 2 (4), 7-41.

Keck, M., & Sikkink, K. (1998). *Activists Beyond Borders: Advocacy Networks in international politics*. Ithaca, NY: Cornell.

Kelly, K. S. (2001). Importance and use of formal research and evaluation. In R. L. Heath (Ed.), *Handbook of public relations* (pp. 279-290). Thousand Oaks, CA: Sage Publications.

Kerstetter, K. (2012). Insider, outsider, or somewhere in between: The impact of researchers' identities on the community-based research process. *Journal of Rural Social Sciences*, 27(2), 99–117.

Lehmann, I. (1999). *Peacekeeping and Public Information, Caught in the crossfire*. London: Frank Cass.

Lehmann, J. (2007). *Evaluating Mass Media Health Communication*. (Unpublished master's dissertation). University of Fribourg, Fribourg, Switzerland.

Lennie, J., & Tacchi, J. (2013). *Evaluating Communication for Development: A Framework for Social Change.* New York, NY: Routledge.

Lewin, K. (1948) 'Group decision and social change', reprinted in M. Gold (Ed.) (1999) *The Complete Social Scientist: A Kurt Lewin Reader.* Washington, DC: American Psychological Association. pp. 265–84.

Lieberman, E. (2005). Nested Analysis as a Mixed-Method Strategy for Comparative Research. *American Political Science Review,* 99(2): 435-452.

Likely, F., & Watson, T. (2013). Measuring the edifice: public relations measurement and evaluation practices over the course of 40 years. In K. Sriramesh, A. Zerfass & J-N. Kim (Eds.), *Public Relations and Communication Management: Current Trends and Emerging Topics* (pp. 143-162). New York, NY: Routledge.

Lincoln, Y. S., & Guba, E. (1985). *Naturalistic inquiry.* London: Sage Publications.

Lindenmann, W. (2003). *Guidelines and Standards for Measuring the Effectiveness of PR Programs and Activities*. Gainesville, FL: Institute for Public Relations. Retrieved from: http://www.instituteforpr.org/topics/effectiveness-programs-activities/

Luo, H. (2010). The Role for an Evaluator: A Fundamental Issue for Evaluation of Education and Social Programs. *International Education Studies*, *3*(2), P42.

Macnamara, J. (2005). Media content analysis: Its uses, benefits and best practice methodology. *Asia Pacific Public Relations Journal*, 6(1), 1–34.

Macnamara, J. (2006). Two-tier evaluation can help corporate communicators gain management support. *PRism*, *4*(2). Retrieved from: http://www.prismjournal.org/fileadmin/Praxis/Files/Journal_Files/Evaluation_Issue/COMMENTARY_MACNAMARA.pdf

Manheim, J. (2011). *Strategy in information and influence campaigns: how policy advocates, social movements, insurgent groups, corporations, governments, and others get what they want*. New York, NY: Routledge.

Mark, M. M., & Henry, G. T. (2004). The mechanisms and outcomes of evaluation influence. *Evaluation*, 10(1), 35-57.

Mason, J. (2002). *Researching your own practice: The discipline of noticing*. London: Routledge.

Mathiason, J. (2011). *The craft of results-based management*. Retrieved from: http://sites.maxwell.syr.edu/intleval/Sessions/1-the_craft_of_evaluation.html

McNiff, J., & Whitehead, J. (2011). *All you need to know about action research.* London: Sage Publications.

Medina-Borja, A., & Triantis, K. (2007). A conceptual framework to evaluate performance of non-profit social service organisations. *International Journal of Technology Management*, 37 (1-2), 147-161.

Mendelsohn, H. (1973). Some reasons why information campaigns can succeed. *The Public Opinion Quarterly*, *37*(1), 50-61.

Meredith, J. (1992). Theory Building through Conceptual Methods. *International Journal of Operations and Production Management,* 13 (5): 3-11.

Michaelson, D., & Macleod, S. (2007). The Application of "Best Practices" In Public Relations Measurement and Evaluation Systems, *Public Relations Journal,* 1(1) 1-14.

Mingst, K. (2004). *Essentials of International Relations*, 3rd ed. New York, NY: W. W. Norton and Company.

Morozov, E. (2009). The Brave New World of Slacktivism, *Foreign Policy*, May 19, 2009. Retrieved from:
http://neteffect.foreignpolicy.com/posts/2009/09/05/from_slacktivism_to_activism?wpisrc=obinsite

Office of the High Commissioner for Human Rights. (2009). *OHCHR   Report: Activities and Results.* Retrieved from:
http://www.ohchr.org/Documents/Press/OHCHR_Report_2008.pdf

Office of the High Commissioner for Human Rights. (2012). *OHCHR Report 2011.* Retrieved from:
http://www2.ohchr.org/english/ohchrreport2011/web_version/ohchr_report2011_web/allegati/3_About_OHCHR.pdf

O'Neil, G. (2013). Evaluation of international and non-governmental organizations' communication activities: A 15 year systematic review. *Public Relations Review,* 39, 572-574.

Otero, E., & Cugelman, B. (2010). *Evaluation of Oxfam GB's Climate Change Campaign.* Oxfam GB.

Patton, M. Q. (2000). Utilization-Focused Evaluation. In D. Stufflebeam, G. Madaus & T. Kellaghan (Eds.), *Evaluation Models: Viewpoints on Educational and Human Services Evaluation* (pp. 425-538). Norwell, MA: Springer.

Patton, M. Q. (2008). *Utilization-focused Evaluation.* London: Sage Publications

Patton, M. Q. (2011). *Developmental evaluation: Applying complexity concepts to enhance innovation and use.* New York, NY: Guilford Pr ess.

Pawson, R., & Tilley, N. (1997). *Realistic Evaluation*, London: Sage Publications.

Pohl, G., & Vanderventer, D. (2001). The workplace, undergraduate education and career preparation. In R.L. Heath (Ed.), *Handbook of public relations* (pp. 357–68). Newbury Park, CA: Sage Publications.

Porto, M. (2007). Fighting AIDS Among Adolescent Women: Effects of a Public Communication Campaign in Brazil, *Journal of Health Communication,* 12, 2, 121-133

Radnor, H. (2001). *Researching your professional practice: Doing interpretive research.* Buckingham: Open University Press.

Reason, P., & Bradbury, H. (2001) *Handbook of Action Research*. London: Sage Publications.

Reeves, T. C. (1997). Established and emerging evaluation paradigms for instructional design. In C. R. Dills & A. J. Romiszowski (Eds.), *Instructional development paradigms* (pp. 163-178), Englewood Cliffs, NJ: Educational Technology Publications.

Rogers, E., & Storey, D. (1987). Communication Campaigns. In C. Berger & H. Chaffee (Eds.), *Reader in Public Opinion and Mass Communication* (3rd ed.) (pp. 419-445). New York, NY: Free Press.

Rogers, P. J. (2008). Using programme theory for complex and complicated programs. *Evaluation*, 14(1): 29-48.

Rossi, P., Lipsey, M., & Freeman, H. (2007). Evaluation: A Systematic Approach. (3rd ed.) Beverly Hills, CA: Sage Publications.

Salmon, T. (1989). (Ed.). *Information campaigns: balancing social values and social chang*e. Newbury Park, CA: Sage Publications.

Salmon, T., & Murray Johnson, L. (2001). Communication Campaign Effectiveness. In R. Rice & C. Atkins (Eds.). *Public communication campaigns. 3rd Edition* (pp. 168-181). Newbury Park, CA: Sage Publications.

Schwandt, T. A. (2003). Back to the rough ground: Beyond theory to practice in evaluation. *Evaluation*, 9 (1), 353–364.

Schwandt, T. A. (2009). Globalizing influences on the western evaluation imaginary. In K.E. Ryan & J.B. Cousins (Eds.), *Sage international handbook on education evaluation* (pp. 119–39). Newbury Park, CA: Sage Publications.

Schwarz, A., & Fritsch, A. (2014). Communicating on Behalf of Global Civil Society: Management and Coordination of Public Relations in International Nongovernmental Organizations. *Journal of Public Relations Research*, 26(2), 161-183.

Scott, R. W. (1995). *Institutions and Organizations*. Newbury Park, CA: Sage Publications.

Scriven, M. (1986). New frontiers of evaluation. *Evaluation Practice*, 7, 7-44.

Shulha, L. M., & Cousins, J. B. (1997). Evaluation use: theory, research, and practice since 1986. *American Journal of Evaluation*, 18(3), 195-208.

Sireau, N. (2009). *Make Poverty History: Political Communication in Action*. Hampshire, UK: Palgrave Macmillan.

Sriramesh, K. (2009). Globalisation and public relations: The past, present, and the future. *PRism* 6(2). Retrieved from:
http://www.prismjournal.org/fileadmin/Praxis/Files/globalPR/SRIRAMESH.pdf

Stake, R.E. (1980). *Program Evaluation, Particularly Responsive Evaluation. Rethinking Educational Research* (pp.72-87). London: Hodder & Stoughton.

Starling, S. (2010). Monitoring and evaluating advocacy: lessons from Oxfam GB's Climate Change campaign, *Development in Practice*, 20 (2), 277-286.

Synder, L. (2007). Health Communication Campaigns and Their Impact on Behavior, *Journal of Nutrition Education and Behavior,* Volume 39, Number 2S, March/April 2007, 32-40.

Tkalac, A., & Pavicic, J. (2009). Nongovernmental organizations and international public relations. In Sriramesh, K., & Verčič, D. (Eds.). *The global public relations handbook: Theory, research, and practice (*2nd ed., pp. 807-821) New York, NY: Routledge.

Tourmen, C. (2009). Evaluators' decision making. *American Journal of Evaluation,* 30(1), 7-30.

Trochim, W.M.K. (2009). Evaluation policy and evaluation practice. In W.M.K. Trochim, M. M. Mark, & L. J. Cooksy (Eds.), *Evaluation policy and evaluation practice. New Directions for Evaluation*, 123, 13–32.

Valente, T. (2001). Evaluating communication campaigns. In R. Rice & C. Atkins (Eds.), *Public communication campaigns* (pp. 105-124). Newbury Park, CA: Sage Publications.

van Ruler, B, Tkalac-Verčič, A., & Verčič D. (Eds.). (2008). *Public relations metrics.* New York, NY: Routledge.

Vogel, I. (2012). *Review of the use of 'Theory of Change' in international development: Review Report.* UK Department for International Development (DFID). Retrieved from: http://r4d.dfid.gov.uk/pdf/outputs/mis_spc/DFID_ToC_Review_VogelV7.pdf

Watson, T. (1995). *Evaluating public relations: Models of measurement for public relations practice.* Article presented to the International Public Relations Research Symposium, Lake Bled, Slovenia.

Watson, T. (1997). Measuring the success rate: Evaluating the PR process and PR programmes. In P. J. Kitchen (Ed.), *Public relations: Principles and practices* (pp. 283–99). Boston, IL: International Thomson Business Press.

Watson, T., & Noble, P. (2007). *Evaluating Public Relations: a best practice guide to public relations planning, research & evaluation* (2nd ed.). London: Kogan Page.

Weiss, C.H. (1998). *Evaluation: Methods for Studying Programs and Policies* (2nd ed.). Newbury Park, CA: Sage Publications.

White, J. (2005). *Measurement and Evaluation: Moving the Debate Forward*. London: Chartered Institute of Public Relations.

Wholey, J. (1994). Assessing the feasibility and likely usefulness of evaluation. In J. Wholey, H. Hatry & K. Newcomer (Eds.), *Handbook of Practical Program Evaluation* (pp. 15-39). San Francisco, CA: Jossey-Bass.

Wholey, J., Hatry, H., & Newcomer, K. (2010). (Eds.). *Handbook of Practical Program Evaluation*. (3rd ed.). San Francisco, CA: Jossey-Bass.

Willetts, P. (2010). *Non-governmental organizations in world politics: the construction of global governance*. London: Routledge.

Williams, B., & Imam, I. (Eds.) (2007). *Systems Concepts in Evaluation: An Expert Anthology*. Point Reyes, CA: Edge Press/American Evaluation Society.

Wilson, L., & Ogden, J. (2008). *Strategic Communications Planning for Effective Public Relations and Marketing (5th Edition)*. New York, NY: Kendall/Hunt Publishing Company.

Woolcock, M. (2009). Toward a plurality of methods in project evaluation: a contextualised approach to understanding impact trajectories and efficacy. *Journal of Development Effectiveness*. Vol. 1(1), 1-14.

Xavier, R., Patel, A., Johnston, K., Watson, T., & Simmons, P. (2005). Using evaluation techniques and performance claims to demonstrate public relations impact: An Australian perspective. *Public Relations Review, 31*(3), 417–24.

Zerfass, A., Verčič, D.,Verhoeven, P., Moreno, A., & Tench, R. (2012). *European Communication Monitor 2012. Challenges and Competencies for Strategic Communication. Results of an Empirical Survey in 42 Countries*. Brussels: EACD/EUPRERA.

Zhao, S., Grasmuck, S., & Martin, J. (2008). Identity construction on Facebook: digital empowerment in anchored relationships. *Computers in Human Behavior,* 2008; 24:1816–36.

Interlude 3

Article 2 provided an analysis and reflection on the methodology used and its implementation in the evaluation of two communication campaigns of IOs. In carrying out the evaluations, this author adopted a pragmatic and adaptive approach, an attempt to find the "middle ground" between the logical and linear evaluation process and the complex activities and contexts of the two campaigns. This produced broad findings on the campaigns and efficiency of their activities but compromised the range of methods that could be used and the consequent strength of the findings. This was tempered by the participative nature of the evaluations that built confidence of organisation staff in its methods and findings.

Article 2 provided further evidence to support article 1 on the factors that influence the ability of organisations to evaluate. The integration of evaluation within the communication function was again highlighted. Article 2 extended this further by identifying other organisational and contextual factors, such as the campaign model and the organisational settings, with a commonality identified being that they were predominantly internal factors.

While article 2 focused on the second component of the evaluation process, Implementation, article 3 focuses on the third and fourth components, Findings and Use. The author returned four years later to the two organisations where the campaign evaluations were carried out to investigate their use. Article 3 extends the proposals of articles 1 and 2 concerning the influence of internal and external factors to these final components of the evaluation process.

6. Article 3: The use of communication evaluation findings in two international organisations

Abstract

This article is an in-depth study on the use of evaluation findings of two communication campaigns of the Office of the High Commissioner for Human Rights and the International Committee of the Red Cross. The research identified 28 instances of use and 6 instances of non-use in these organisations with the large majority being unanticipated and instrumental in nature. Use decreased when moving away from the campaign managers both in time and in distance. The strongest influences on use were found to be internal; relating to organisational context and communication goals and ambitions, which could enable or impede use, depending upon the situation. Evaluation use travelled from the individual to the organisational level in a predominantly non-linear fashion, interconnected, overlapping and bringing about change both in a formal and informal manner but never in a vacuum devoid of other influences. Use was found to be unpredictable and constructed by the meaning assigned by staff members, adjusting and interpreting findings in opportunistic and unexpected ways.

Keywords: public relations measurement, campaigns, communication evaluation, evaluation use, evaluation utilization, international organisations, non-profit communications, evaluation methodology.

6.1. Introduction

Within the evaluation community there is a rare consensus that a key purpose of evaluation is that its findings should be used for improving programmes and decision-making (Alkin & Taut, 2002; Henry & Mark, 2003; Patton, 2008).   As a consequence, it follows that evaluation use or utilization has received arguably the most attention and research of the evaluation field with considerable empirical reviews and conceptual studies carried out (Alkin, Daillak, & White, 1979; Christie, 2007; Ciarlo, 1981; Cousins, Goh, Elliott & Bourgeois, 2014; Johnson et al., 2009; Patton et al., 1977; Weiss, 1979). Yet, it has been found that contemporary theories of evaluation use are "simultaneously impoverished and overgrow" (Mark & Henry, 2004, p. 37); "Impoverished" in that there has been little understanding of the underlying processes that lead to use; "Overgrown" in that too much attention has been paid to the conceptualisation and categorisation of use.

The contribution of this article is to the "impoverished" side of evaluation use research; an in-depth study on the use of evaluation findings in two international organisations (IOs) with a focus on understanding the ways individuals and organisations use evaluation findings and the interrelations between the evaluation process, people, organisations and contexts.

This article investigates the use of findings from the evaluation of two communication campaigns of two IOs, the Office of the High Commissioner for Human Rights (OHCHR) and the International Committee of the Red Cross (ICRC). These evaluations were carried out by one of the authors of this article from 2009 to 2010 and some four years later, the authors returned to study their use. In examining instances of use and their path through the organisation and interrelation with different factors, this article differs from the existing body of research which largely emphasises the role of the evaluation process, methods and products on use (Cousins and Leithwood, 1986; Leviton & Hughes, 1981). More so, the article makes reference to recent theoretical contributions on the change processes that mediate evaluation use and the role of factors beyond the evaluation process   while extending this further to consider reception and meaning theories

(Contandriopoulos & Brousselle, 2012; Cousins, Goh, Clark & Lee, 2004; Fitzpatrick, 2012; Henry & Mark, 2003; Højlund, 2014a; Mark & Henry, 2004; Patton, 2000).

Simply put, the interest of this article is less on *what type of use occurred? How did the evaluation product facilitate use?* And more so on *how did use happen? What is the process of use? What enables and impedes use?*

The definition of evaluation use provided by Johnson et al. (2009) guides this article: "any application of evaluation processes, products, or findings to produce an effect" (p.378). The concept of use has been expanded by the notion of evaluation influence, defined as the "capacity or power of persons or things to produce effects on others by intangible or indirect means'' (Kirkhart, 2000: p. 7). This article is more orientated towards the definition of *use* rather than *influence* and examines more the *direct* effect of evaluation findings *within* the organisations ("use"), rather than their indirect or intangible influence.

6.2. Models of evaluation use

6.2.1. Existing models of evaluation use

Interest in evaluation use dates from the 1960s when there was concern that evaluations were being carried out with little regard as to how their findings would and were being used (Alkin & Taut, 2002; Ciarlo, 1981). This led to a focus on understanding how evaluation use could be increased through a series of studies and theoretical reflections that are considered as cornerstones of the field (Alkin, Daillak, & White, 1979; Ciarlo, 1981; Patton et al., 1977; Weiss, 1979). Over time, typologies of use were developed, debated and validated by empirical studies with most centring on four types of use: instrumental, conceptual, symbolic and process, with the latter being a more recent addition (Ciarlo, 1981; Cousins & Leithwood, 1986; Shulha & Cousins, 1997; Patton, 2000). Instrumental use refers to instances where knowledge from an evaluation has been used directly. Conceptual use refers to instances where people's understanding has been affected but no direct action has taken place. Symbolic use refers to instances

where the evaluation was used a basis for action (or inaction) or to justify pre-existing positions. Process use refers to instances where the involvement in the evaluation process provoked changes (Alkin & Taut, 2002; Johnson et al., 2009).

Studies of use have showed limited evidence of direct (instrumental) use but more so conceptual, symbolic and process use (Ciarlo, 1981; Cousins et al., 2004; Patton et al., 1977). Research also focused on determining the factors that influenced use or not, with a commonly cited set of 12 factors as those used by Cousins and Leithwood (1986). Six factors concerning the evaluation implementation have been identified: evaluation quality, credibility, relevance, communication quality, findings and timeliness; and six factors identified on the decision or policy setting: Information needs, decision characteristics, political climate, competing information, personal characteristics and receptiveness/commitment. Cousins and Leithwood found the most influential factors were evaluation quality followed by decision-making characteristics. Using a similar framework, Højlund (2014b) found that decision and policy setting were considerably more important than evaluation implementation factors.

These four types of use and 12 factors of influence constitute the first widely used model of evaluation use. Although the totality of the model and these empirical and conceptual studies are recognised as significant in shaping this field of inquiry (Shulha & Cousins, 1997), there has been criticism of their limitations. As a response, the field has been broadened in recent years, with the model of evaluation use adapted and reframed into a second generation of theories and models.

One major criticism was that the preliminary model emphasised the evaluation process and products as the key influence on use at the expense of other factors, notably human and context (Alkin & Taut, 2002; Højlund, 2014a). The importance of context has been cited since modern day evaluation commenced yet it has rarely been considered in-depth (Fitzpatrick, 2012). Contandriopoulos and Brousselle (2012) found that attention has been concentrated on the means and ability of the evaluator to influence and encourage use, while not recognising that context might be the "essential determinant" (p. 71).

Different factors of context began to take further prominence from 1980s onwards, including resource scarcity (Mowbray, 1992), organisational structures and processes, programme aspects (Mathison, 1994; Torres, Preskill & Piotnek, 1996), institutional contexts (Højlund, 2014a) and the evaluation system within organisations (Højlund, 2014b).

 The rise of context as a factor was also linked to a broadening of the epistemological base of evaluation; the realist and positivist approaches underlying evaluation were increasingly challenged by critical theory, constructivist and interpretive approaches (Albaek, 1995; Shulha & Cousins, 1997). A concrete consequence was the rise in the practice of participatory evaluation together with an expansion of the role of the evaluator; from that of an external "judge" or "investigator" with more positivist interpretations to that of "facilitator", "problem-solver", "coach" or "critical friend" with more constructivist interpretations (Caracelli, 2000; Scriven; 1986).

 As a consequence, in their empirical review of evaluation use from 1986 to 2005, Johnson et al. (2009) added an additional category of "stakeholder engagement" to that of Cousins and Leithwood's 1986 framework and model. Their review concluded that "engagement, interaction, and communication between evaluation clients and evaluators is key to maximizing the use of the evaluation in the long run" (p. 377). This confirmed earlier findings of Shulha and Cousins (1997) citing Ayers (1987), Cousins (1995) and Greene (1988) who emphasised the link between use and participation, notably that stakeholders' participation gives them confidence to use research procedures, assurance of the quality of evaluation findings and a sense of ownership in the findings and their consequent application.

The human factor as an influence on evaluation has long been recognised in the literature although the orientation has been towards the importance of the personal characteristics, means and ability of the evaluator rather than the intended user (Contandriopoulos & Brousselle, 2012; Cousins & Leithwood, 1986; Johnson et al., 2009; Shulha & Cousins, 1997). Cousins and Leithwood's 1986 framework did include factors concerning the

intended user, such as their role in the organisation, receptiveness and information needs. However, aside from assessing the relative influence of these factors on evaluation use, there has been limited perspective offered on the reception process of users, that is, how does a user receive, interpret, filter and decode evaluation findings.

One significant development in this second generation of theories and models has been the work of Henry and Mark (2003 & 2004). They identified and conceptualised the underlying processes through which evaluation can exercise influence. Drawing from organisational, social and behaviour theories, they extended the model of evaluation use with a set of mediators and pathways for change at the individual, intrapersonal and collective organisation levels. The significance of this development lay in its attempt to unpack how use and influence can occur at different levels in organisational settings in addition to confirming the multidimensional nature of evaluation use. There have been two known and documented attempts to apply this adapted model to actual studies of use, but in both cases they failed citing difficulties to adjust their research methodologies and collect the data needed for pathway modelling (Johnson et al., 2009; Weiss, Murphy-Graham & Birkeland, 2005).

6.2.2. Communication evaluation and use

In the area of communication and campaign evaluation, little attention has been paid to how findings are used, although there is an implicit assumption in the literature and field of the importance and intention of use (Broom & Dozier, 1990; Stacks & Michaelson, 2010; Watson & Noble; 2007). In a study of communication evaluation reports of IOs, 96% considered improvement to programmes as seen in their findings and recommendations (O'Neil, 2013); Gregory and White (2008) point out the contribution of communication to improved decision-making, organisational functioning and ultimately organisational value and how evaluation can support this.

The predominant theoretical models of organisational communication are rooted in systems theory and foresee a role for communication managers to help organisations adapt to their environment, providing inputs from their environment, one of which would

be evaluation findings (Broom & Dozier, 1990; Dozier, Grunig & Grunig, 2001). Further the most progressive of these, the two-way symmetrical model, was based on the dialogical concept with the use of feedback from multiple sources, including evaluation for the purpose of organisational adaptation (Broom & Dozier, 1990).

Although communication evaluation models have been criticised for their absence of links to organisations, their environment and processes, Watson and Noble (2007) emphasise the formative role of evaluation by providing feedback to the organisation with the aim of increasing the effectiveness of future campaigns and other communication activities. Recent communication evaluation models that are more participatory, such as those proposed for development settings draw on complexity and systems theory with the stated aim of increasing utilisation of evaluation findings (Lennie and Tacchi, 2013). Yet, caution is sounded by Zerfass (2008) who found that many evaluation activities for communications are "mere rituals of verification" (p. 150). Further, that many organisations are "closed", impervious to external inputs through communications, and that the dominant communication model in practice is asymmetrical with more information flowing out of the organisation than into it (Broom & Dozier, 1990; Dozier et al., 2001).

6.2.3. Towards the next generation of models of evaluation use

This second generation of theories and models on evaluation use have drawn from a broad range of areas including organisational theory, psychology, political science, social and behavior theory, learning theory and practice (Alkin & Taut, 2002; Cousins et al., 2004; Henry & Mark, 2003; Højlund, 2014a; Kirkhart, 2000).

However, theories and models have been limited in considering the comparative nature of evaluation as a concept, the knowledge it produces and the reception process of evaluation findings.

Integrating further the perspective of systems theory, organisational processes and decision-making would provide a better understanding of how learning and change occur

in organisations and in relation to individuals; how decisions are taken, policies created and what factors of influence are at play, including and in addition to, evaluation findings (Birkland, 2014; Katz & Kahn, 1966; Saaty, 1999; Sutton, 1999). For example, this perspective would consider evaluation as essentially an external concept introduced into an organisation which comes with its own field, systems, institution, procedures, expertise and is concretely operationalised through evaluation studies and their anticipated use. Here there are parallels to be drawn from the experience of introducing other external concepts into organisations. For example, Mannell (2014) describes studies that have been made on introducing the concept and policies of gender within non-governmental organisations and the identification of supporting factors for its implementation, resembling studies made on evaluation and its use. Cohen and Levinthal (1990) developed the concept of "absorptive capacity", the extent to which organisations recognise the value of new, external concepts and information and use it innovatively, challenging the "closed" organisation described above. Research is yet to compare evaluation and its use to the experience of introducing other external concepts within organisations where studies exist, whether it is cross-cutting fields such as gender and environmental impact or more applied concepts such as performance measurement or knowledge management (Bhatt, 2001; Hall, 2008; Julnes & Holzer 2001; Levy 1992).

Theories and models of evaluation use often reflect policy formation and decision-making in organisations as relatively linear and rational processes (Alkin & Taut, 2002; Ciarlo, 1981; Cousins et al., 2004; Henry & Mark, 2003). However, this is at odds with the policy science literature which describes the process of taking decisions and creating policies as "evolutionary", "non-rational", "messy" and even "chaos" (Clay & Schaffer, 1986; Juma & Clark, 1995; Sutton, 1999). In this regard, when considering the influence of evaluation findings on policy and decision-making, the potential unpredictability of these processes could be taken further into account.

Concerning the reception process of users of evaluation findings, the work of de Certeau (1990) serves as a useful reference. De Certeau distinguished between the production of policies and their enforcement ("strategies") and how they are actually used by people

("tactics"). He argued that these strategies will never be completely followed but nor will they be overturned by their users; more so users through their actions (tactics) will interpret policies (strategies) differently and in ways that are opportunistic and not anticipated (what he labelled "bricolage"). Parallels can also be drawn to broader debates in evaluation use. There is an underlying assumption in the literature that the evaluation and its findings are "right"; that users should be "instrumentalised" in their implementation of findings and consequently research has focused on what elements (factors of influence) will improve this. De Certeau argued that the user actually has more power to influence their use of strategies (evaluation findings in this case) than thought of; users could give the pretext of implementing while in fact resisting and constantly adjusting their use, which is not quite "misuse", as envisaged in the literature (Caracelli, 2000). Studies of evaluation use would therefore be astute in considering the tactical side of use (de Certeau's terminology), that is the reception process and meanings created.

The notion of meaning, that is, how do people in the reception process interpret and understand has been reinforced by Luhmann (1990) in his theory of the "improbability of communication". He highlighted three obstacles to communication occurring; 1) meaning, the extent to which one person can understand what another meant is highly dependent upon the individual and context; 2) reach; that it is improbable that more people than present in a given space and time will be reached; and 3) acceptance, that even if the two previous obstacles are overcome, there is no guarantee that what is being communicated will be accepted. More so, it will be processed with other experiences, thoughts and perceptions to construct meaning. This theory is of interest when examining evaluation use to consider how people interpret differently the same findings; is it possible to reach people beyond the immediate recipients of the evaluation report, and how will people accept a given finding and balance it with other experiences and information?

The significance of meaning has been understudied in evaluation use. Its importance was emphasised in a recent study where 19 different contextually-driven interpretations were found of one simple phrase ("I have to go to work"), none of them pertaining to its

standard dictionary definition (Gillespie & Cornish, 2014). In this regard, theories and models of evaluation use could be reinforced through considering more in-depth the reception process of users.

Considering the above, this article seeks to make a contribution to the modelling of evaluation use with the intention of developing a more comprehensive model. As seen in Figure I below, this recognises the evolutionary nature of the model, that is, that existing models are built on rather than abandoned. The first and second generation of models represents visually their main aspects as described above. The main development from the first to the second model is the recognition of context and human factors and the addition of levels of use (Mark & Henry, 2004).

In the proposed next generation model, evaluation is recognised as a continual rather than a linear process (e.g. "process to findings"). The accompanying circle represents the organisational setting where the evaluation is received and meaning created from which evaluation use emerges. Non-use is added as a second type of use, meaning that an evaluation finding could be considered but then not taken up or rejected. Categorisations are purposely left general as use or non-use may occur, anticipated or not. The different levels of use are represented in a non-linear fashion. Factors of influence are grouped as "enabling" and "impeding" recognising that a given factor may operate as both depending upon the circumstances. Further, these factors should be recognised as potentially influencing both the evaluation process and its use (O'Neil, 2015).
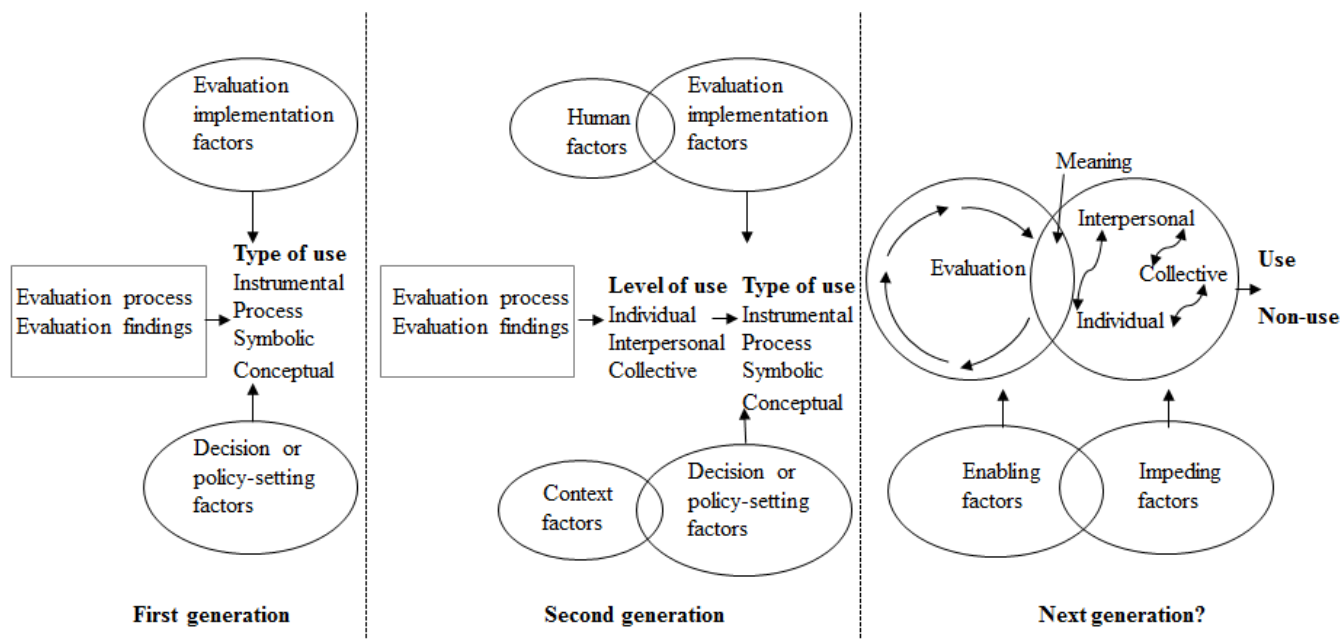
Figure I: Evolution of the model for evaluation use

Through examining the use of findings from the evaluation of two communication campaigns, this article will contribute to understanding the applicability of this next generation model.

6.3. Method: case studies of OHCHR and ICRC

This article investigates the use of findings from the evaluation of two communication campaigns of the OHCHR and the ICRC. The starting point for the study was interviews with the campaign managers who commissioned and managed the evaluations in 2009 (OHCHR) and 2009-2010 (ICRC). The managers were asked for other staff to interview within the communication department that knew of and potentially used the evaluation findings; staff interviewed were asked the same (snowball sampling technique). This led to six interviews with ICRC staff and five interviews with OHCHR staff.   Interviews were semi-structured and focused on exploring instances of use from the campaign evaluations and/or validating instances mentioned by other staff. To analyze the data, a conceptual framework was created with five dimensions: Type of use based on the commonly-used typology described above; level of use and attributes of change both based on the framework developed by Henry & Mark (2003); influences of use was taken

from the conceptual framework described below; and instance validation based on strategies used in previous studies (Ciarlo, 1981; Højlund, 2014b; Weiss et al., 2005). Questions of reception and meaning were discussed during the interviews. The factors for each dimension are detailed in Table I. Each instance of use was coded on the basis of this framework. The coded results of instances of use are detailed in the table found at the end of this article. Each instance of use is given a label for easy reference (e.g. ICRC5).

Table I: Conceptual framework for analysis of use

Type of use	*Level of use*	*Attributes of change*	*Influence on use*	*Validation*
• Conceptual • Instrumental • Non-use • Process • Symbolic	• Individual • Interpersonal • Collective	*Individual* • Attitude change • Behaviour change • Elaboration • Priming Knowledge acquisition* • Salience • Skills acquisition *Interpersonal* • Change agent Consensus* • Exchange* • Justification • Minority-opinion influence • Persuasion • Social norms *Collective* • Agenda setting • Diffusion • Policy	*Internal* • Communication goals and ambitions • Evaluation policies and institutions • Organisational context *External* • External context • Field *Internal/external* • Funding • People	• Documentation • Other Staff

		change		
		Policy-oriented learning		
		• Practice change*		

These attributes were added by the authors during the analysis.

In addition to the above analysis, a second methodological element used was the reflection and experience of the authors in carrying out the two campaign evaluations and its implications for use. For this purpose, the "insider-researcher" approach was drawn upon (Mason, 2002; Radnor, 2001) although the role was closer to what has been labelled as the "insider-outsider" approach; that is the "space between" insider and outsider researchers (Dwyer & Buckle, 2009; Kerstetter, 2012). The authors were outsiders as external researchers but at the same time insiders given that the one of the authors had previously worked closely with the organisations in the evaluations.

The third methodological element used was a conceptual framework for communication evaluation for IOs as seen in Figure II. Developed by one of the authors, this framework sets out the evaluation process in relation to internal and external factors that are considered to have an influence on an organisation's ability to carry out communication evaluation. The framework draws from systems theory, organisational behaviour and existing empirical and theoretical research in the communication and evaluation fields (Bamberger, Rugh & Mabry, 2011; Christie & Fierro, 2012; Dubin, 1976; Fitzpatrick, 2012; Højlund, 2014a; Manheim, 2011; Meredith, 1992; Scott, 1995; Trochim, 2009; Watson & Noble, 2007).
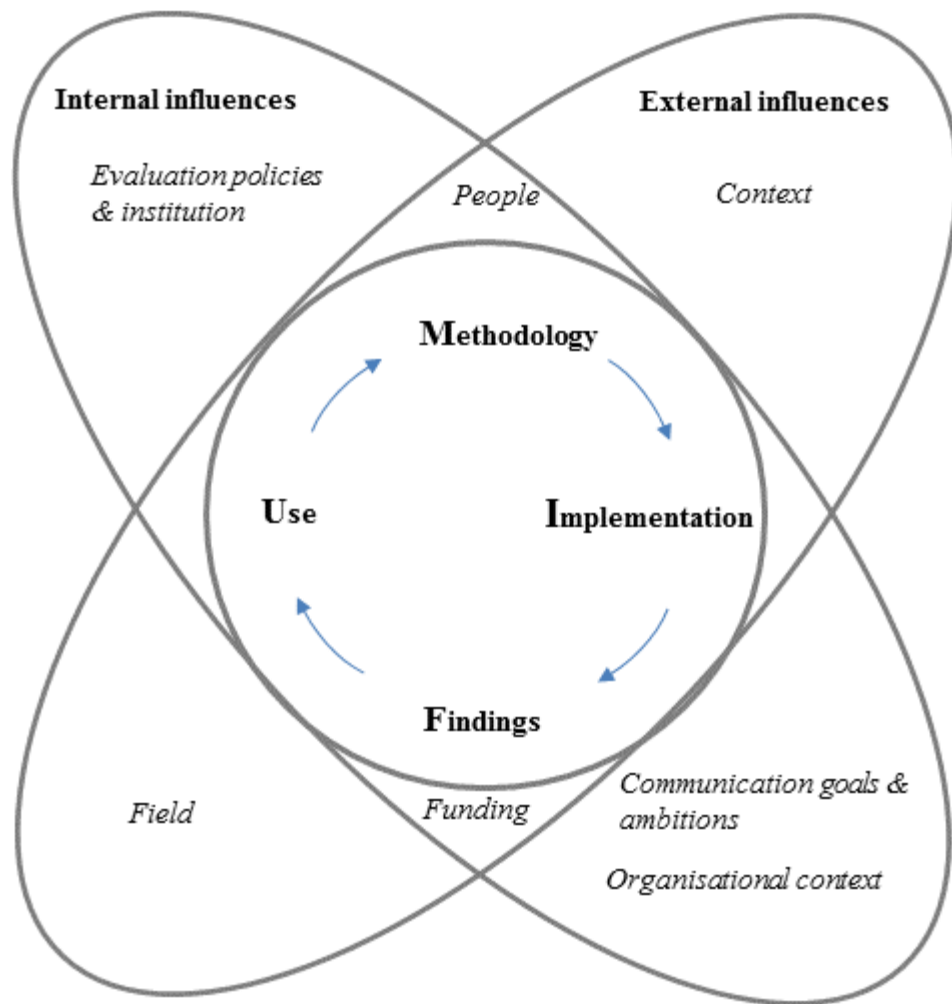
Figure II: MIFU conceptual framework for communication evaluation for IOs

The framework highlights the interaction between the four components of the communication evaluation process (Methodology, Implementation, Findings and Use) and the possible influence from factors outside but interlinked to this process. These factors are described in Table II below, based on the above-mentioned references and used in this article as one of the above-described dimensions of analysis. The focus of this article is on the third and fourth components of the evaluation process, "Findings" and "Use".

Table II: Internal and external factors of the conceptual framework	
Evaluation policies and institutions	Internal to the organisation, evaluation policies, guidelines and direction as managed by a central evaluation unit; the main components of the larger evaluation system.
Organisational context	Internal contextual elements such as structure, culture and strategy.
Communication goals and ambitions	The scope of the communication activities being evaluated, such as the level of effects being sought and the implementation models used.
Context	The external setting within which an organisation carries out its communication activities.
Field	The community of organisations and bodies that have a common meaning system (e.g. peer organisations, academia and industry associations). Two fields are relevant for this study: evaluation and communication evaluation.
Funding	The financial resources available to communication evaluation.
People	The competencies of the human resources implicated in communication evaluation (e.g. communication staff, evaluation staff and external evaluation consultants).

6.4. Results: use and non-use

The research identified 15 instances of use and 3 instances of non-use in the ICRC and 13 instances of use and 3 instances of non-use in the OHCHR. Validation of instances was possible through either documentation and/or confirmation of other staff. A detailed list and categorisation of the 34 instances (as per Table I) is found in Table IV, located at the end of this article. Figures III and IV provide a visualisation for each organisation of the identified instances of use:

-The sources of instances, staff members, are found in the circles to the left;

-The instances of use, categorised by type of use, are found in the rectangles in the centre;

-The influences on use are found in the diamonds to the right;

- The lines indicate the connections between these elements; thickness of the lines represents the strength of the connection (also summarised in numerical values found next to the circles, rectangles and diamonds).

There is some duplication in these calculations as staff members (sources) often cited the same instances of use. For example, in figure III, there are seven instances of instrumental use from 12 sources, as five instances were mentioned by more than one source.
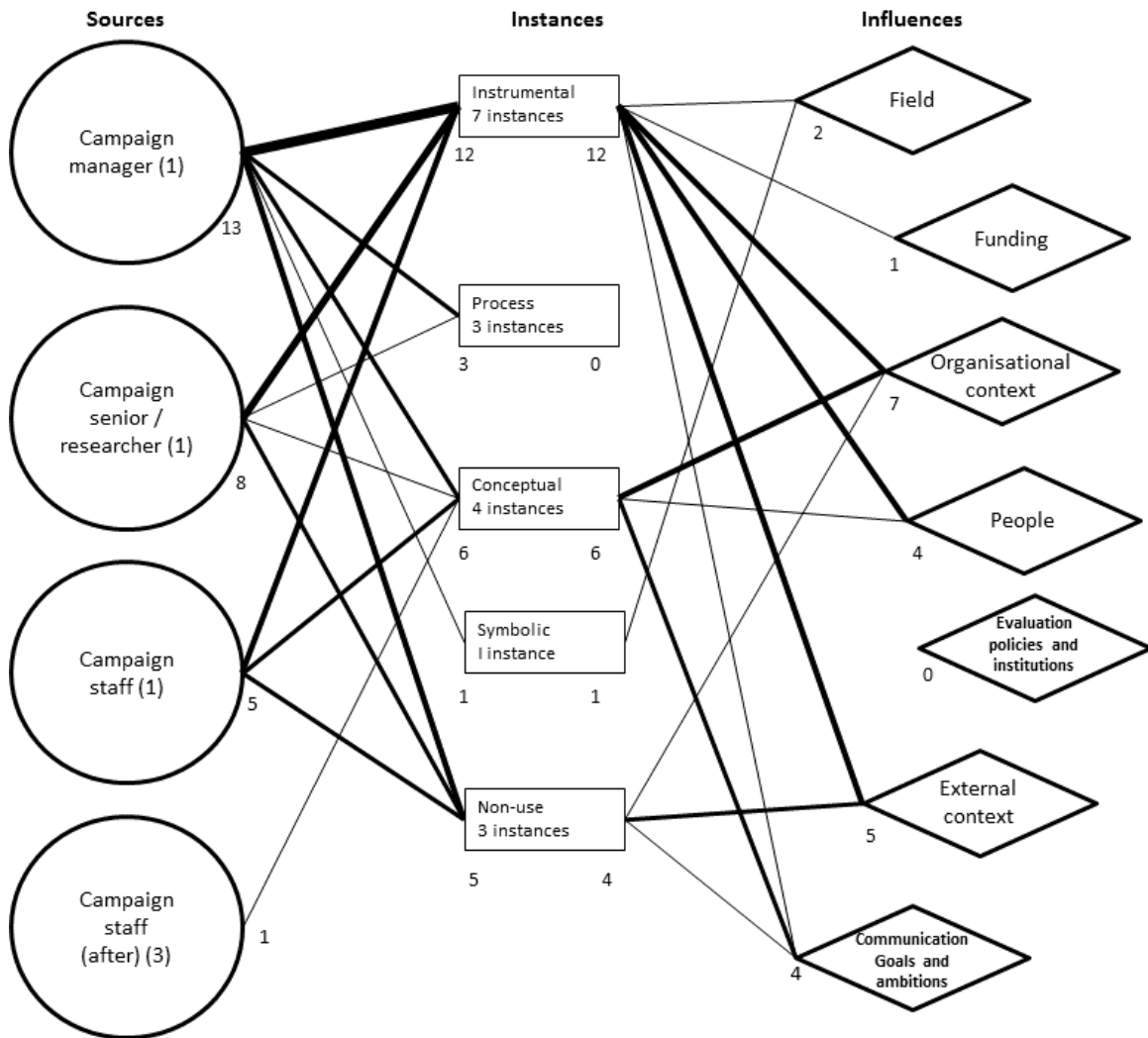
Figure III: Instances of sources, types of use and influences – ICRC
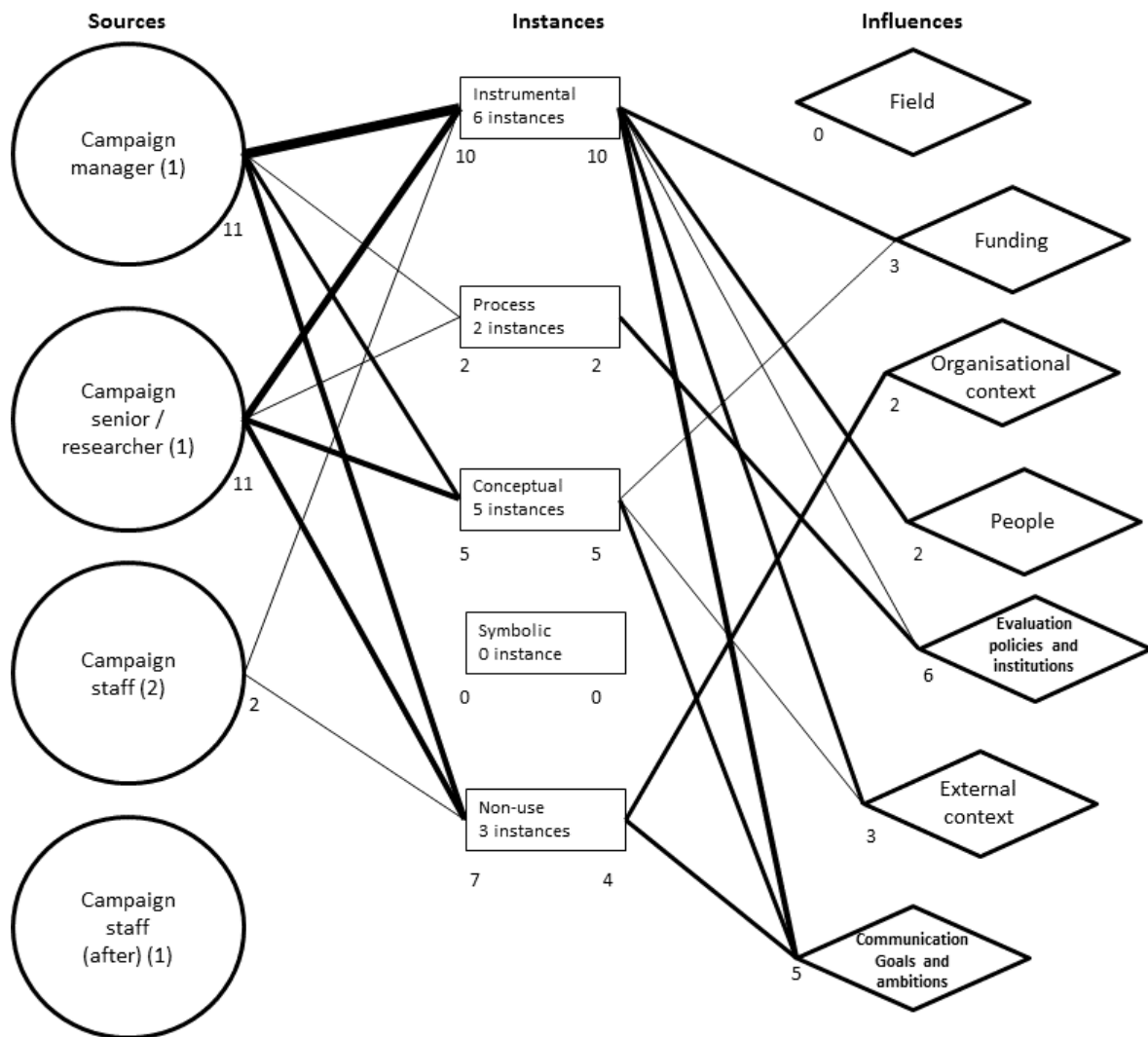
Figure IV: Instances of sources, types of use and influences – OHCHR

6.4.1. Reception and meaning

The reception process differed notably by the staff members' roles. Campaign managers in both organisations commissioned the evaluations and were involved in all steps of the evaluation process (O'Neil, 2015). Therefore, they viewed the evaluations as a direct feedback on the campaigns they managed and were those who could provide the most instances of use (24 in total) which were predominantly instrumental as can be seen in Figures III and IV. The senior campaign/research staff members of both organisations were similarly involved in the evaluation process and could cite 19 instances of use. Less

involved in the evaluations were campaign staff who could provide fewer instances of use (7 in total). Only one campaign staff who started after the evaluation (at the ICRC) was able to cite an instance of use, symbolic in nature (ICRC14).

Consequently, meaning drawn from the evaluation findings varied. Of note, out of the 34 instances identified, 16 were stated recommendations of the evaluation findings and 18 were not, that is, they were "unanticipated". Further, if the six instances of non-use, all stated recommendations, are discounted, the number of instances of unanticipated use is nearly double in comparison (18 compared to 10). For unanticipated use, these were instances which were drawn from the findings text, implied or explicit but were not recommendations (what Kikhart (2000) referred to as "unintended" use). These instances, mostly identified by campaign managers, were their own construction of meaning drawn from the evaluation and in some cases, there was even a sharing of meaning between staff on what was not explicitly stated in the evaluation findings. For example, for instance OHCHR1, both the campaign manager and senior campaign staff member identified "review timing of campaign material" spontaneously as the first instance of use they thought of, although it was not a stated recommendation or explicitly advocated in the report.

Out of the four staff that started after the evaluation, only one ICRC staff member knew of the evaluation and could cite an instance of their use of the findings (ICRC14). More so, these staff could cite instances of campaign policy or practices that had changed over time but they did not know that the evaluation findings had influenced these changes. For example OHCHR interviewee (P4) mentioned "importance of partners has increased as has our support", which was an instance of use cited by two other colleagues. This meant that any significant meaning assigned to the evaluations was relatively lost for these staff not directly involved in the evaluations, but consumed into a pool of generalised knowledge, similar to what Weiss (1981) observed:

> They [program managers and decision makers] have absorbed the generalizations from diverse sources over a period of time, and these ideas become the taken-for-granted assumptions on which they base new plans and decisions. (p. 23)

At the same time, the meaning assigned to the totality of the evaluations differed in the organisations. Within the ICRC, the evaluation was the last major evaluation of a communication campaign and had taken on a type of symbolic status amongst the campaign manager and staff members, as ICRC interviewee I2 commented "whenever I read a new concept note for a campaign or communications action I share the [evaluation] report with the relevant staff; it became a reference point for me". In comparison, within the OHCHR, since the completion of the evaluation, the organisation has carried out four more evaluations of campaigns (as their campaigns were run annually). Therefore in the interviews with the campaign manager and staff, feedback on the evaluation overlapped and was mixed with that of the other evaluations with often a distinction blurring as seen when OHCHR interviewee O1 discussed process use: "I learnt from this evaluation but also from those that followed, it all works together in that way".

There was no evidence that the evaluation findings had travelled further beyond the limited circles of campaign staff, although a limitation of the study was that not all possible users of the findings were known or possible to identify. As seen above, use was concentrated on those that had direct interaction with the evaluations. However, this did not limit the possible wider influence of the evaluation findings on broader policies as seen in several instances of use, that is, where the campaign manager acted upon a finding that then impacted on such policies (e.g. ICRC11, OHCHR10).

In both organisations, instrumental use was the dominant type of use found (7-ICRC; 6-OHCHR); conceptual use followed (4-ICRC; 5-OHCHR) and it increased relatively for campaign senior/researcher role compared to campaign managers. It was also the most durable over time, i.e. it was the only type of use cited by a staff member who started after the evaluation.   For both organisations, campaign managers and campaign senior/research staff members were the only staff that could cite instances of process use (3 – ICRC; 2 - OHCHR) which is understandable as these were the staff that were heavily involved in the evaluation. Instances of non-use were equally cited by ICRC and OHCHR

campaign managers and campaign senior/researcher staff members (3 – ICRC; 3 - OHCHR). One instance of symbolic use was identified by the ICRC campaign manager.

6.4.2. Influences on use

Influences on use were discussed for each instance with the relevant staff. For the ICRC, as illustrated in Figure III, the strongest influence on use was organisational context followed by external context, communication goals and ambitions and people. For OHCHR, as illustrated in Figure IV, the strongest influence on use was evaluation policies and institutions followed by communication goals and ambitions. Evaluation policies and institutions was the weakest influence for the ICRC as was field for OHCHR.

It was found that most influences could operate both as "enabling" or "impeding" factors depending upon the given instance (Højlund, 2014b; Mark & Henry, 2004). For example for organisational context, in the instance ICRC2 "Greater involvement of Red Cross Movement in campaigns", this finding was very much in line with the new organisational strategy for 2015-2018 that encouraged closer work with the Movement. In an instance of non-use (OHCHR15) "Finding a balance between NGO and UN compatible messages", the communication team were constrained by organisational priorities in their selection of messages. For communication goals and ambitions, the campaign model used by both organisations integrated strongly the role of partners which meant the organisations were not directly managing all campaign aspects, limiting their ability to implement evaluation findings on monitoring or in-country implementation (ICRC18, OHCHR14). At the same time, the campaign model facilitated those findings that emphasised further integration of partners (OHCHR3, OHCHR4, ICRC12).

External context was influential notably in the changes in the communication landscape with the move towards digital channels and therefore facilitated any evaluation findings in this direction (OHCHR5, 6 & 11) and discounted those that did not fully consider the changing landscape (ICRC17). Simultaneously, the complex and varied environments where the organisations sought to communicate impeded implementation of evaluation findings (ICRC5 & 18).

For the OHCHR, evaluation policies and institution was an enabling factor by encouraging (and in some cases requiring) planning, monitoring and evaluation supporting findings in this area (OHCHR2, 7, 8, 9, 10 & 12); this was not the case for the ICRC as discussed further below. For both organisations, people as an influence was an enabling factor, in that the views of the communication staff coincided with a number of the evaluation findings or recommendations, facilitating their implementation (ICRC1, 3, 4 & 12; OHCHR3 & 4). Funding was influential in four instances (ICRC1, OHCHR3, 4 & 10) where availability of funding limited the ability of the organisations to implement all aspects or delayed them. The influence of field was only seen in two instances, both with the ICRC (ICRC4 & 15). This was where the organisation re-considered the use of certain campaign tactics based on the evaluation findings, which was also supplemented by consultation with peer organisations (part of "Field") on their use of tactics.

By type of instances, instrumental use was influenced mainly by people, the external context and communication goals and ambitions. Influences on process use were limited to evaluation policies and institution, in that what was being learnt through the evaluation process was supported by this factor. Conceptual use was less influenced and strongest being organisational context and evaluation policies and institutions. Non-use was influenced by three factors, organisational context, external context and communication goals and ambitions.

6.4.3. Pathways of use

Henry & Marks (2003) pathways of influence was modelled on the notion of evaluation findings bringing influence to bear as they travel through three levels within organisations – from individual to interpersonal to the collective - in a type of causal chain. In some respects, this study found this model applicable. For example, all the instances of use had their origins at the individual level. Communication staff individually reflected on the evaluations findings and considered their implications before they were rejected or advanced to the interpersonal and eventually the collective level, with a diminishing number of instances advancing from step to step. There was no

instance found where use had its origins at the interpersonal or collective level. But the pathways model was also found to be a simplification of how influence and decision-making actually happened in the organisations studied, where use did not always occur in a strict linear and rational fashion.

Communication staff in both organisations described how the evaluation findings and their use took place in a series of interrelated and complex processes. Information from the evaluation and other sources was digested by individual communication staff, balancing this with enabling or impeding influences and existing beliefs and information. Discussions with colleagues were held to seek out their opinions and find a consensus. Decision-making often worked in a cyclical manner jumping back and forth between the individual and interpersonal levels or even skipping levels (i.e. interpersonal). A culmination of efforts could lead to change(s) being made (or not), formally to policies and informally to practices to manage campaigns and communication in general. Scenarios of this nature worked in parallel, interconnected and overlapped with some accelerating quickly and concluding within months whereas others could continue for many years. This description supports the literature that challenges the rational theory of organisational behaviour and decision-making; that decisions are rarely taken rationally and in a single, isolated moment but more so are dealt with in multiple discussions and meetings (Clay & Schaffer, 1986; Juma & Clark, 1995; Mintzberg & Waters, 1985; Sutton, 1999).

However, in some cases, use or non-use did occur in a direct and linear fashion. An example being instance ICRC4, where the evaluation report recommended focusing on campaign activities that "work well" in global campaigning and de-prioritising those that did not. The campaign manager in consultation with her staff implemented this in the next campaign strategy through the selection of campaign activities. Enabling influences also supported this; the recommendation coincided with the views of the campaign manager and staff ("people") and sharing of experiences with peer organisations ("field").

On this basis, each instance could be categorised on the basis of how use or non-use

occurred, if it was anticipated or not; linear or not. This can be organised into a process categorisation as described in the following table and labelled using travel analogies:

Table III: Process categorisation of how use occurred in both organisations			
Label	Description	Number of instances	
		ICRC	OHCHR
1. Direct route	Use was anticipated and occurred in a linear way.	4	1
2.Unexpected hop	Use was not anticipated and occurred in a linear way.	5	4
3. A planned ramble	Use was anticipated and occurred in a non-linear way.	2	3
4. Unforeseen foray	Use was not anticipated and occurred in a non-linear way.	4	5
5. Expedition starts/stops	Use was anticipated, did not occur and happened in a non-linear way.	1	3
6. Surprise trip deferred	Use was not anticipated, did not occur and happened in a non-linear way.	0	0
7. Travel plans cancelled	Use was anticipated, did not occur and happened in a linear way.	2	0
8. Unannounced stop-over skipped	Use was not anticipated, did not occur and happened in a linear way.	0	0

As seen in Table III, use/non-use occurring in a non-linear way (18 – categories 3, 4, 5 & 6) was slightly more predominant than use/non-use in a linear way (16 – categories 1, 2, 7 & 8). According to the evaluation use literature, the most expected way that use would occur would be category 1: Anticipated linear use. However, this was not a dominant way that use occurred; more so it was unanticipated and could be equally linear or non-linear (categories 2 and 4). An example and graphic illustration is provided of each category in the following paragraphs.
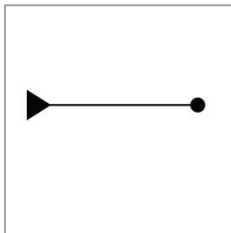
1. *Direct route:*   Instance ICRC1 (Reduce complexity of messages and products) is an example where a recommendation of the evaluation report was taken up by the communication manager, agreement found with the team and implemented in the next campaign launched. Implementation was facilitated given that the

recommendation confirmed the beliefs of the campaign staff and manager and was believed to lead to possible cost-savings.
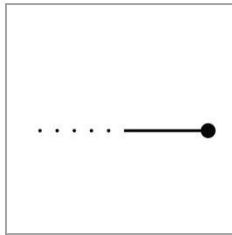
2. *Unexpected hop:* Instance OHCHR8 (Learnt monitoring and evaluation terminology), through participating in the evaluation, a communication staff member learnt directly of monitoring and evaluation terminology, such as "outcomes" and "indicators". This learning was also influenced by the introduction of the RBM system that was occurring at the same time of the evaluation and used similar terminology.
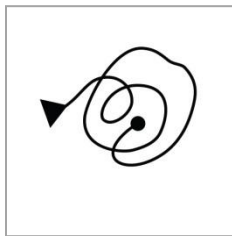
3. *A planned ramble*: Figure V shows the instance OHCHR4, an example of anticipated non-linear use. The non-linear nature of the use is evident as is the impeding and enabling influences and the eventual informal and formal changes seen at the collective level.
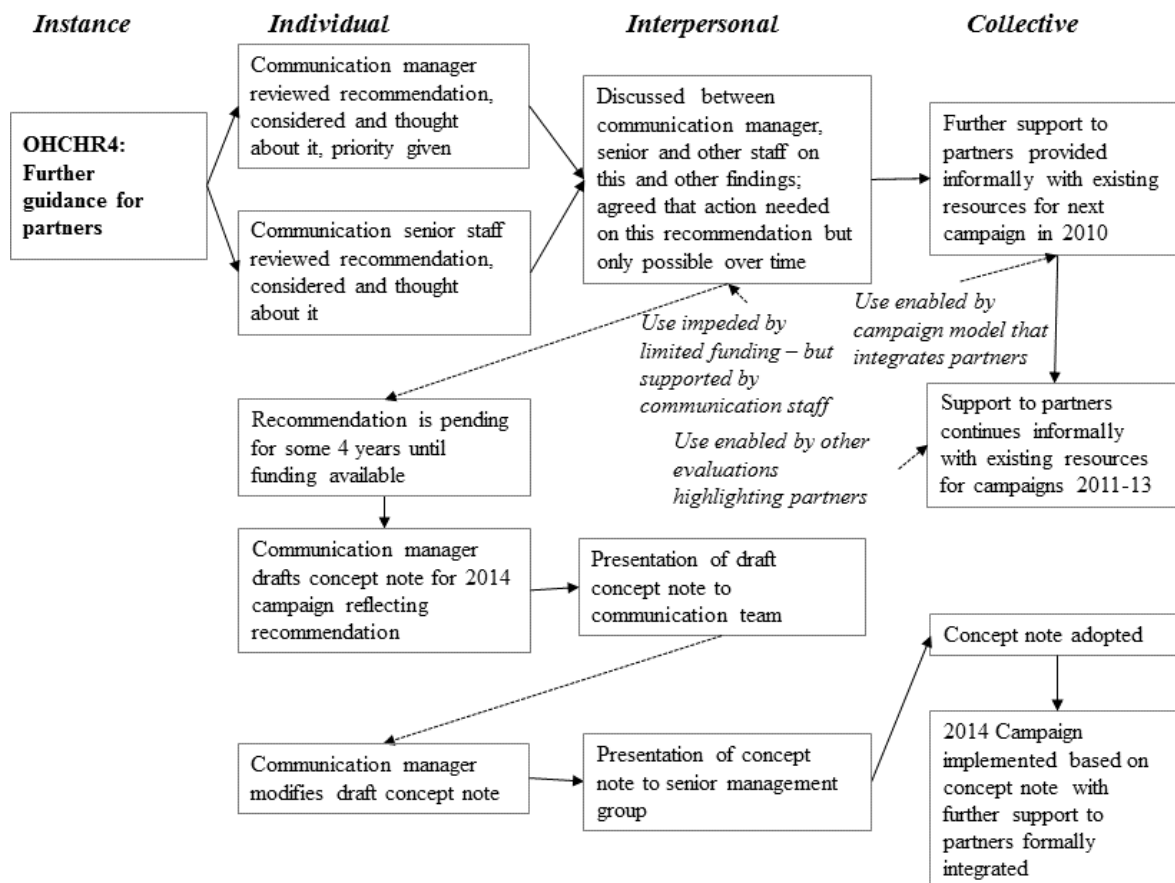
Figure V: Example of anticipated non- linear use –instance OHCHR4
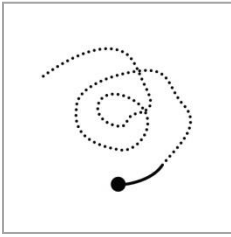
4. *Unforeseen foray:* Instance ICRC11 (possibility to use campaigns to mobilise publics) is an example where the evaluation report provided an unanticipated input into internal discussions on the role of communication and mobilising publics that were ongoing at the time of the evaluation and continued for the following years. This was considered non-linear as described by staff, discussions were happening in parallel and moving between the individual to interpersonal level (and back again), with an anticipation that policy revision would occur. The evaluation was one of the many inputs into the policy revision.
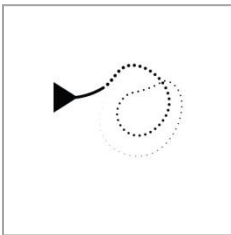
5. *Expedition starts/stops:* Instance OHCHR14 (set measurable objectives and target audience) was a specific recommendation of the evaluation report that was considered by the campaign manager and not directly dismissed. The campaign manager discussed this with his staff on several occasions before concluding it was not feasible, given that the campaign model relied largely on partners which complicated efforts to be more specific and precise in targeting and measurement.
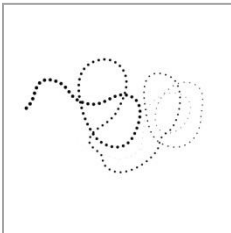
6. *Surprise trip deferred:* There were no examples found of this type. This scenario would be where an organisation has considered an aspect of evaluation findings that was not explicitly stated as a recommendation, reflected on it over time in multiple and circular discussions and reflections, and finally decided not to take any action.
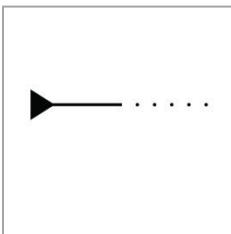
7. *Travel plans cancelled*: Instance ICRC17 (integrate further needs of low technology contexts) was a specific recommendation of the evaluation report that was considered by the campaign manager and dismissed directly without internal discussion or reflection. In taking this decision, the campaign manager indicated that rapidly changing contexts where the organisation was working was the main influence; the gap between

low and high technology contexts (e.g. access to internet and mobile networks) was narrowing rapidly and would be further so by the next planned major campaign.
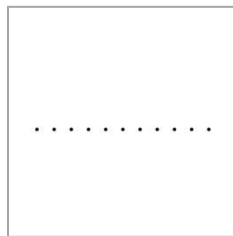
8. *Unannounced stop-over skipped*: There were no examples found of this type. This scenario would be where an organisation has considered directly an aspect of evaluation findings that was not explicitly stated as a recommendation and dismissed it immediately.

To understand further the interactions within and between the three levels, a simplification of the pathways and compilation of instances from both organisations of use/non-use was mapped out, as illustrated in Figure VI. For the sake of the analysis, this simplification essentially eliminates the linear and non-linear element and does not show all the connections, back-and-forth movements and links as shown in the above example (Figure V).
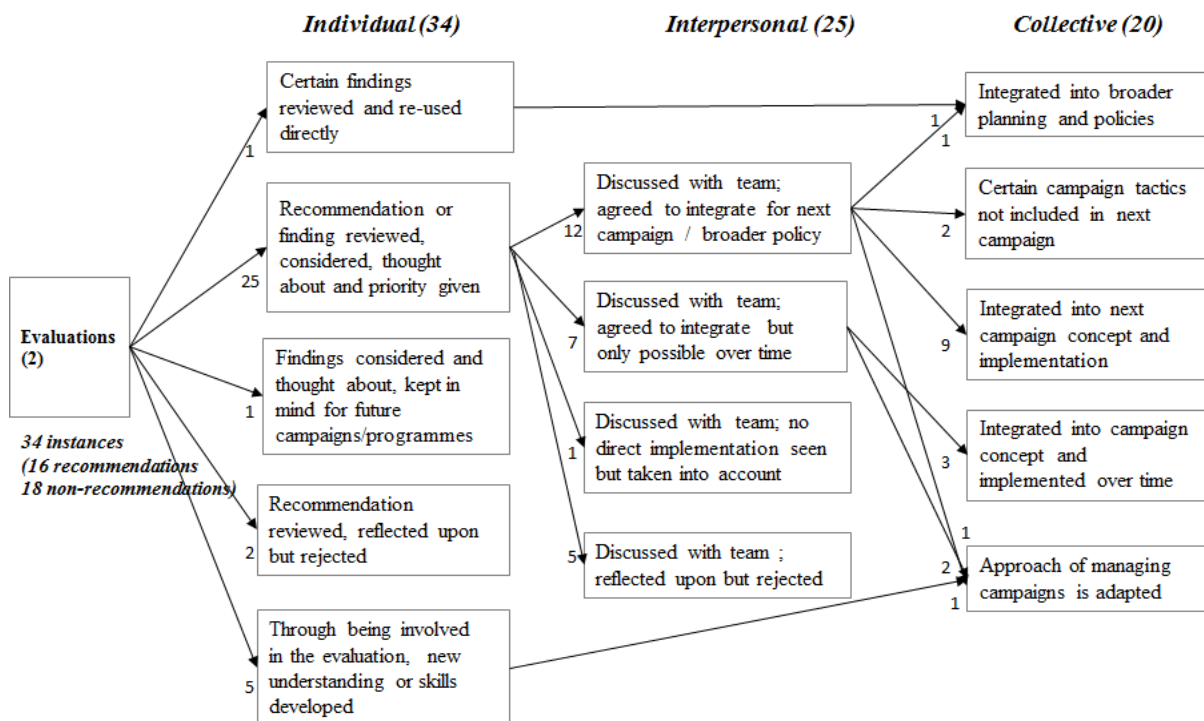
Figure VI: Simplified pathway of use through individual, interpersonal and collective levels

At the individual level, recommendations and findings of the evaluation reports were reviewed and considered, mostly by the campaign managers and senior staff/researchers. On this basis, 25 instances, mostly instrumental in nature, were given priority for discussion with other communication staff (at the interpersonal level). Some findings (notably stated recommendations) were rejected (2 instances of non-use) or kept in mind for the future but no action taken (1 instance). By being involved in the evaluation, new understanding or skills were developed by individuals (5 instances of process use) with one instance leading to influencing the approach or practice of managing campaigns.

The interpersonal level was conceived as change being brought about by the interaction between individuals prompted by evaluation findings. In the two organisations, 25 instances of use were identified at this level, all stemming from the individual level. The setting for the interpersonal level was a discussion on an evaluation finding or findings, either informally between communication staff or more formally in a meeting, for example if it involved staff from other units. The given finding, highlighted mainly by the campaign manager was discussed and debated. Some half (12 instances) were agreed to be integrated for the next campaign and others were agreed to be integrated but only over time (7 instances) – thus these advanced to the collective step either directly or after several discussions and reflections at both the individual and interpersonal levels. One instance was taken into account but no direct implementation was seen and five instances were reflected upon and then rejected. These latter instances were findings or recommendations where the campaign manager wanted to discuss them with colleagues before rejecting them, and were not continued with further, mostly because of internal influences, such as communication goals and ambitions and organisational context; and external context to a lesser extent.

The collective level was conceived as where change would occur at the organisational level, for example, where a policy change was influenced by the evaluation findings. In the two organisations, 20 instances of use were identified at this level, with 18 stemming from the interpersonal level and two from the individual level. The setting at the collective level was commonly the integration of evaluation findings into policy and

practices, for example, in campaign concepts (9 instances) that guided the next campaigns to be implemented. In three instances, they were not integrated immediately but into future campaigns. Two instances were also seen where it contributed to the organisations not to continue with a given campaign tactic. Two instances were seen which inputted into broader communication policies and planning. Four instances influenced more the approach of how campaigns were managed, such as prioritization to a given area of work (e.g. monitoring or setting objectives), which these authors label as "practice change" (Mark, Cooksy & Trochim, 2009).

6.4.4. Attributes of change processes

The typology of change processes that could be triggered by an evaluation at each of the three levels developed by Henry & Mark (2003) was applied to each instance of use.

At the individual level, the evaluation findings went through a process of reception, selection and meaning as described above. In most instances identified, they were then a trigger for raising the given issue in the minds of the communication staff (attribute of salience), that led them to thinking it over and developing it further (elaboration). In some cases, this in turn influenced their opinion or strength of their opinions in bringing it to the 'top of their mind" (priming). By being involved in the evaluation process, one instance of skills acquisition was found but it was mostly knowledge that was acquired, an addition to the original typology of Henry and Mark. In reality, the evaluation findings triggered these attributes but they overlapped and were interconnected with other influences and existing beliefs of staff. For example, an evaluation finding could trigger more thought on the specific issue in an individual, bringing it to the forefront of many competing issues, while simultaneously reinforcing their existing opinion and providing momentum to raise the issue with colleagues. In this regard, the thought process as described by communication staff was compatible with the cognitive psychology concept of mental models (Johnson-Laird, 1983; Vandenbosch & Higgins; 1996). That is, staff integrated the evaluation findings into their existing beliefs that then supported their actions and decisions, mostly confirming existing models in these two cases rather than

creating new ones (Hall, 2011; Vandenbosch and Higgins, 1995; Vandenbosch and Higgins, 1996).

The major variations to the original typology were seen at the interpersonal level. Whereas Henry and Mark (2003) describe a setting where individuals seek to influence others through persuasion and as agents of change, what was found in these two organisations was a more subtle process of discussion and consensus building. This is supported by the literature on group dynamics and organisational change that emphasizes the development of shared understanding and reconciliation of conflicting perspective that happens largely through group discussions and processes (Mohammed & Ringseis, 2001; Sutton, 1999). Issues brought by individuals to the interpersonal level were discussed and in most cases a consensus found as to whether to proceed or not to the collective level, with discussions and reflections back and forth between these two levels. This did involve persuasion and justification in some instances to convince colleagues of the worth of the issue but was also compounded by colleagues with supportive or non-supportive opinions (influenced or not by the evaluation) and other sources, such as personal experiences or available monitoring data. In turn this led to a setting where interaction did bring about change but more often there was common agreement found rather than individuals competing to influence.

At the collective level, the typology of Henry & Mark (2003) was more extensive than what was seen within the two organisations studied. Where agreement was found with the existing typology was that the evaluation findings did trigger policy change in some cases, albeit not always immediately and influenced by other sources as described above. Further, an additional attribute was added of "practice change", an informal change to the way that a campaign was managed, such as the prioritization to a given area of work (e.g. monitoring or setting objectives). Practice normally has no written dictate and has been referred to as "informal implicit rules" (Mark, Cooksy & Trochim, 2009, p.6) and is still considered as a type of policy. A number of the attributes anticipated by the typology were not found within the identified instances of use, such as agenda setting (moving of issue on the public agenda) and diffusion (influence on another sector or jurisdiction).

This is also due to the fact that the evaluations were internal to the organisations, within a specific technical area (communication) and not foreseen to influence a larger policy debate, as was the case in the examples cited within the original typology of Henry and Mark.

6.4.5. Similarities and differences in use between the ICRC and OHCHR

Where similarities were found between the organisations was in the types of use (predominantly instrumental and conceptual use), sources of use (mainly campaign managers and senior/research staff members), the pathway through the organisations (gradual decrease of use from individual to the collective) and the predominance of non-linear and unanticipated instances over linear and anticipated instances. Differences were seen in two aspects. Firstly, some differences were seen with the influences of use, with the greatest being the factor of evaluation policies and institutions. This factor was identified as influencing six instances of use within OHCHR whereas none within the ICRC. OHCHR staff referred to the emphasis placed on planning and evaluation by its relatively new performance monitoring system and thus saw this as a key influence on use. ICRC staff referred to a similar consciousness of planning and evaluation but did not name its source. An explanation could be that their equivalent to a performance monitoring system had been in place for over a decade at the time of the evaluation and was thought of as a given (O'Neil, 2015). Secondly, the meaning attached to the evaluation in its totality differed between the two organisations with the evaluation having a more symbolic status in the ICRC whereas in OHCHR it was consumed within other evaluations, as discussed earlier.

The communication function of the ICRC was larger than that of the OHCHR in terms of staff and resources, which meant that potentially the evaluation findings could have had greater use across the staff. However, this was not found to be the case, as the total instances of use/non-use of the two organisations were relatively similar (OHCHR-16, ICRC - 18) although not all instances of use may have been discovered by this study.

6.5. Discussion: use constructed by meaning, setting and context

This study found that the staff members directly involved with the evaluations were central to their use with instances originating mostly with campaign managers that would then travel to the interpersonal and in some instances to the collective levels. The notions, guidance or recommendations from the evaluations that crystallised as instances of use were not all found to be predictable but rather constructed by the meaning given by staff members. In the two organisations studied, use was found to never take place in a vacuum, it competed with other sources and was mainly influenced by internal factors. Of these factors, organisational context and the communication goals and ambitions were found to bear the strongest influence, enabling and impeding, depending upon the situation. The contribution of evaluation findings to change within individuals and organisations has to be seen in the setting where decision-making and policy construction was interconnected, overlapping and more cyclical than linear in nature, with changes occurring both in a formal and informal manner.

This study showed an interesting phenomenon at work concerning the meaning assigned to the evaluations. This was evident in that over half of the instances of use found were unexpected and not explicit in the evaluation findings. In this way, it confirmed de Certeau's theory (1990) of "bricolage" that the campaign manager and staff interpreted the evaluation findings in opportunistic and unanticipated ways and Luhamann's theory on "improbability of communication" (1990) and its three obstacles. Suggested by de Certeau, these unexpected notions or guidance drawn from the evaluation findings did not go completely against the evaluations, but fitted within their overall direction. In this regard, the staff effectively resisted any imposition of the evaluations but also went further by rejecting recommendations that they felt not appropriate or incompatible with their organisational and external contexts, communication goals and ambitions. Kirkhart (2000) identified such unexpected use as being of interest given that its scope could go far beyond the intended and classic users of the evaluations. In these two cases, overall meaning assigned to an evaluation was also found to be dependent upon the frequency of

evaluation occurring; evaluation carried out regularly meant that their findings and consequent use was fused and joined, losing some distinction of individual evaluations; evaluation carried out infrequently took on more of a symbolic and distinct meaning.

The experience taken from these two organisations indicates that it seems near impossible to predict which evaluation findings will resonate the strongest within organisations. In addition, the study showed that distance and time impacted on getting the attention of potential users of the evaluation (Luhmann, 1990) with a limited circle of communication staff reached. Further, an organisation with a larger communication team (ICRC) did not necessarily mean that use was greater than in a smaller team (OHCHR).

Previous studies in evaluation use have found limited evidence of instrumental use. This study found the contrary; use was predominantly instrumental in nature. This could be explained by the particular context of the evaluations. Those who were the greatest potential users of the evaluation, the campaign managers and their direct staff, directly commissioned the evaluations and were heavily involved in the process from shaping the methodology to data collection and inputting into findings (O'Neil, 2015). This confirms the link between participation and use as emphasised in the literature (Ayers, 1987; Braverman & Arnold, 2008; Cousins, 1995; Greene, 1988; Johnson et al., 2009; Shulha & Cousins; 1997). This participatory approach adopted for the evaluations and the consequent interaction between the evaluator and the staff was key to maximising use, as illustrated by a comment of an ICRC interviewee I2 "we always had a dialogue with the evaluator and felt we also owned the findings".

The topics of the instances that went on to contribute to policy change show some similarities. An examination of these nine instances (Figure VI and Table IV) shows ideas or recommendations that focus on improving the efficiency of campaign relays, tactics and messages or emphasise how campaigns could be better used to support organisational goals. This is of interest as the main focus of both evaluations was on evaluating the outputs/outcomes of the campaigns. However, interaction with staff members during the evaluations led to adjustment of the evaluation goals to take into

account staff members' interest in this respect (O'Neil, 2015). As a consequence, what campaign staff reacted to and took action on was mainly in improving the efficiency and relevance of future campaigns, and by inference, their own performance and that of their teams. This also indicates that the new information and perspectives that evaluation provided did not necessarily lead to the creation of new mental models for staff but more so confirmed and strengthened existing models (Johnson-Laird, 1983; Vandenbosch & Higgins; 1996).

The strongest influences on evaluation use were largely contextual, but internal rather than external, namely the organisational context and the communication goals and ambitions. This is a key distinction considering that these internal factors are in theory more possible to control and manage compared to external contextual factors, such as changes in the communication landscape. As seen in the above analysis, such internal contextual influences could be both enabling and impeding dependent upon the instance of use in question. Although the human factor (people) was found as a moderate influence, it should be recognised that the communication staff were key to use by placing the given evaluation finding in the appropriate discussion or process and advancing it further. This corresponds to the findings of Weiss (1998) that evaluation findings bring new information that competes and sometimes integrates with the existing three I's of staff members; ideology, interest and information. The enabling influence of the evaluation policies and institutions in the case of OHCHR supports previous studies that showed the positive influence on use when evaluation is institutionalized (Højlund, 2014b).

Studying these influences should also not be limited to only the use of evaluation findings. More so their study can be extended to the whole evaluation process, as described in Figure II. This is possible in studies such as this one that was preceded by a study of the evaluations' implementation (O'Neil, 2015). Of interest, both this and the preceding study both indicated that internal factors of influence were the most important influence on the evaluations, in their conceptualisation, implementation and use.

The two cases of this article found that the contribution of evaluation findings to change within individuals and organisations has to be seen in the complex and interconnected processes in which findings are digested and integrated. Although the linear three-level model of Henry and Mark facilitated analysis, decision-making and policy construction was found to be far more complex. The evidence showed evaluation findings could contribute to change at the policy level. However, it was not always as expected, largely tempered by other influences and could be both formal and informal, such as changes to practices.

These findings illustrate that use will not always happen immediately and directly; and that over time, it is possible that certain aspects of the findings will become engrained in the policies and practices and their origin will be lost and absorbed into an organisational body of knowledge of multiple sources, as previously proposed in the literature (Kirkhart, 2000; Weiss, 1981; Weiss; 1998). What is not fully explained by this study is the relative influence of evaluation and its findings against other competing or complementary sources and factors at the policy level.

6.5.1. Limitations

A limitation identified in carrying out this study was that one of the authors had previously carried out the evaluations and then returned to assess their use, bringing into questions of potential bias. Studies of evaluation use have previously faced this same situation (Ciarlo, 1981; Højlund, 2014b; Russ-Eft, Atwood & Egherman, 2002) but none addressed the issue of bias explicitly. For this article, these authors sought to minimise bias by involving a second author and using a validation strategy for any instance of use claimed as described above.

Another limitation identified was the question of time delay and recall of evaluation use, given the 4-5 years between the evaluation and this study of use. Alkins (2002) suggested that the optimal time period for use was one year. This study didn't find issues with recall; on the contrary, the extended period worked in its favour in identifying longer term examples of use.

6.6. Conclusion: valid and useful contribution of evaluation

This study is a contribution to the body of evaluation use research and has focused on understanding the way individuals and organisations use evaluation findings and the interrelations between the evaluation process, people, organisations and contexts. This study confirms the relevance of systems theory to evaluation use, that is, the action of bringing in an external element such as evaluation findings cannot be viewed in isolation but has to be considered in relation to other influences and the interlinking and relation to other processes that contribute to change.

Based on this study, the authors agree with Contandriopoulos & Brousselle (2012) that context can be an essential determinant in evaluation use but would nuance this to *internal* context as seen in the two organisations studied. Although this study has focused on the influence on factors other than the evaluation process, this study confirmed the finding of Johnson et al. (2009) that a participatory evaluation approach did facilitate use. Previous studies have showed a low prevalence of communication evaluation in companies and organisations (30-50%) and even lower in IOs (13%) (Macnamara, 2006; O'Neil, 2013; Watson, 1997; Xavier, Patel, Johnston, Watson & Simmons, 2005). Yet, this study showed that evaluations can be of use and are used by communication professionals. In the two cases studied, findings were used mainly to improve efficiency of future campaigns, even if it was done in areas unexpected and in a somewhat opportunistic way, that is, staff extracting meaning from findings that mostly supported their interests and priorities. Nevertheless, it indicates that evaluation as a concept and practice brings a valid and useful contribution to communication professionals.

Returning to the contribution of this article to the next generation of evaluation use models and theory, the findings illustrate the strong linkages between the evaluation process (i.e. staff participation) and evaluation use. The importance of meaning constructed from use and the existence of factors of influence was found to both impede and enable usage, influencing not only use but the whole evaluation process. Both linear

and non-linear paths to change were also found to produce use, anticipated and unanticipated.

This study examined the use of findings of evaluation reports, a punctual activity. Given that monitoring of communication and campaign activities has been documented as another important element of evaluation (Starling, 2010), in future research it would be of interest to study its use and influence.

Finally, if evaluation use research is to focus on understanding its relative influence on individuals and organisations then it may need to go about it in another way. For example, instead of studying the use of evaluation findings as this and other studies have done, an alternative way would be to study a policy in question and assess the relative influences of evaluation and other factors on its creation. Further, research could also move further away from the "overgrown" area on conceptualisation and categorisation and as suggested earlier, consider studies of comparison of evaluation to other external concepts introduced into organisations.

Table IV: Coding and description of instances of use

#	So.	An.	Instance description	Type	Individual		Interpersonal		Collective				
					Description	Att.	Description	Att.	Description	Att.	How	Inf.	Ver.
ICRC 1	I1, 2, 3	y	Reduce complexity of messages and products	In	Recomm. reviewed, considered, thought about, priority given	Sa, El	Discussed with campaign team, confirmed existing consensus	Ex, Pe	Integrated into next campaign concept and implemented in messages/produ cts developed	POC	ALU	F, P	D; I1, 2, 3
ICRC 2	I1, 3	y	Greater involvement of Red Cross Movement in campaigns	In	Recomm. reviewed, considered, thought about; priority given	At, Sa, El	Discussed with team, considered important for some future campaigns	Ex, Co	Integrated into campaign planning & general approach	POC, PRC	ANLU	OC	D; I1, 3
ICRC 3	I1, 3	y	Include a central "ask" in future campaigns	In	Recomm. reviewed, considered, thought about, priority given	Sa, El	Considered, decision made to include in next campaign	Ex, Ju, Co	Fundraising central "ask" included in next campaign	POC	ALU	P, C	D;I1, 3
ICRC 4	I1, 2	y	Focus on activities that "work well"; de-prioritise those that do not	In	Recomm. reviewed, considered, thought about, priority given	At, Sa, El	Discussed with team & complemented by own reflections	Ex, Pe	Certain activities integrated in next campaign; others not	POC (2)	ALU	P, FI	D; I1, 3
ICRC 5	I1	y	Consider option of using campaigns as operational tools	In	Recomm. reviewed, considered, thought about	Sa, El	Recomm. aligned with team reflection & adopted for next campaign	Ex, Co	integrated into next campaign concept	POC	ALU	OC, C	D; I2
ICRC 6	I2	n	Learning that quantitative	In	Finding reflected upon,	Sa, El	Discussed with team; agreed for	Ex, Co	Integrated into next campaign	POC	ULU	CG	D; I2

			results attracted media coverage		consideration for future		next campaign strategy		concept and implemented in campaign				
ICRC 7	I2	n	Action to share report with colleagues who plan and manage campaigns	In	Concepts and strategies of other programmes reflected upon	Sa	Feedback provided to managers	Pe	Integrated into relevant concept of campaign	POC	ULU	OC, C	I2, 3
ICRC 8	I1	n	Understand what evaluation methods can be used (and reused)	Pr	Developed understanding of evaluation methods & use	At, Sa, El	None	n/a	None	n/a	ULU	n/a	I1
ICRC 9	I1	n	Understand the limits of evaluation	Pr	Developed understanding of the limits of evaluation for campaigning	At, Sa, El	None	n/a	None	n/a	ULU	n/a	I1
ICRC 10	I2	n	Learn about campaign strategies and tactics	Pr	Learnt about campaign strategies and tactics during the evaluation	Sa	None	n/a	None	n/a	ULU	n/a	I2, 3
ICRC 11	I1, 3	n	Possibility to use campaigns to mobilise publics	Con	Findings on public mobilisation reviewed and contributed to reflection	At, Sa, El	None	Ex, Ju	Inputted into planned revisions of policy on communications	POC	UNLU	OC, CG	I1, 3
ICRC 12	I1, 2	n	Greater involvement of field offices in campaign	Con	Findings on field offices reviewed and contributed to	At, Sa, El	Taken into consideration in discussions on policy	Ex, Ju	Inputted into planning for future campaigns	POC	UNLU	OC, CG, P	I2, 3

			strategy and design		reflection		modifications concerning public mobilisation						
ICRC 13	I1	y	Consider the different needs of field offices	Con	Recomm. reviewed, considered and thought about	Sa, El	Taken into consideration in discussions for future strategy and design for campaigns	Ex, Co	None	n/a	ANLU	OC	I1
ICRC 14	I4	n	Importance of developing clear objectives and measuring results	Con	Findings considered and thought about, kept in mind for future programmes	At, Sa, El	Recomm. aligned with own reflection but no direct implementation seen	Ex, Co	None	n/a	UNLU	n/a	I4
ICRC 15	I1	n	Dismissing proposed tactics that the evaluation indicated were unsuccessful	Sy	Finding on non-successful tactics recalled	Sa, El	none	Ex, Ju	Certain campaign tactics not included in campaign strategy.	POC	UNLU	F	I1
ICRC 16	I1	y	Consider ability to reach non-traditional audiences	N-U	Recomm. reviewed, taken into consideration but no direct action taken; not precise or actionable	Sa, El	Evaluation findings cited in discussions with colleagues on potential campaign tactics	n/a	None	n/a	ALNU	n/a	I1
ICRC 17	I1, 2	y	Integrate further needs of low-tech contexts	N-U	Recomm. reviewed, reflected upon but rejected as obsolete given rapidly changing	Sa, El	None	n/a	None	n/a	ALNU	C	I1, I2

					environment								
ICRC 18	I1, 2	y	Consider the different needs of National Societies	N-U	Recomm. reviewed, considered and thought about	Sa, El	None	Ex, Ju	None	n/a	ANLNU	C, OC, CG	I1, I2
OHCHR 1	O1, 2	n	Review timing of campaign material	In	Findings reflected upon, consideration for future programmes	Sa, El	Discussed with team but not implemented as issue difficult to respond to given campaign model	Ex, Pe, Co	Integrated into approach by campaign team for future campaigns	PRC	UNLU	CG	D; O1, O2
OHCHR 2	O1, 3	n	Use of quantitative findings for annual planning process	In	Select findings reviewed & re-used directly	Sa, El	Discussed with team, agreed that action needed	n/a	Data of evaluation report used for annual planning process	POC	ULU	E	D; O1, O3
OHCHR 3	O1, 2	y	Harnessing partners worldwide	In	Recomm. reviewed, considered and thought about	Sa, El, Pr	None	Ex, Ju, Co	Level of integration of partners is adapted over time	POC	ANLU	F, P, CG	D; O1, O2
OHCHR 4	O1, 2	y	Further guidance for partners	In	Recomm. reviewed, considered, thought about, priority given	Sa, El	Discussed with team, agreed that action needed to be taken but only possible over time	Ex, Ju, Co	Process of supporting partners adapted over time	POC	ANLU	F, P, CG	D; O1, O2
OHCHR 5	O1	n	Modify priority of communication tactics	In	Findings reflected upon, consideration for future programmes	Sa, El, Pr	Discussed with team, agreed that action needed to be taken but only possible over time	Ex, Ju, Co	For future campaigns, certain tactics maintained others reinforced	POC	ULU	C	O1

OHCHR 6	O2	y	Incorporating public mobilisation in campaigns	In	Recomm. reviewed, considered and thought about	Sa, El, Pr	Discussed with team, justifies previous action (choice of some tactics) and need to adjust on others	Ex, Pe	Tactic is tested in future campaign	POC	ALU	C	D; O2
OHCHR 7	O1	n	Skills and know-how on campaign design	Pr	Learnt about campaign design during the evaluation	El; Sa	Discussed with team, agreed to test tactic in next campaign	n/a	The way of designing campaigns is adapted	PRC	UNLU	E	D; O1
OHCHR 8	O2	n	Learnt monitoring and evaluation terminology	Pr	Learnt of monitoring & evaluation terminology	At, Sa, El	None	n/a	None	n/a	ULU	E	O2
OHCHR 9	O1	n	Priority of monitoring and evaluation for staff	Con	Findings prompted reflection on role of staff	Sa, El	None	Ex, Co	None	n/a	UNLU	E	O1
OHCHR 10	O1	y	Alternative measures for evaluating awareness	Con	Recomm. reviewed, considered & thought about, given priority	Sa, El	Discussed with team but no immediate action taken	Ex, Ju, Co	More emphasis given to monitoring & evaluation	PRC	ANLU	E, F	D; O1
OHCHR 11	O2	n	Increase use of social media and web for campaigns	Con	Importance of issue raised; thought about and given priority	Sa, El, Pr	Discussed with team, agreed that action needed to be taken but only possible over time	Ex, Ju, Co	None	n/a	UNLU	C	D; O2
OHCHR 12	O2	n	Further forward planning for campaigning	Con	Importance of issue raised; thought about and given priority	Sa, El, Pr	Discussed with team; no direct action taken, more so overall approach to campaigning	Ex, Ju	None	n/a	ULU	E	O2

OHCHR 13	O2	n	Understanding theoretical aspects of campaigning	Con	Importance of issue raised; thought about and given priority	Sa, El, Pr	Discussed with team; no direct action taken, more so approach over next years	Ex, Ju	None	n/a	UNLU	n/a	O2
OHCHR 14	O1, 2, 3	y	Set measurable objectives and target audiences	N-U	Recomm. reviewed, considered and thought about	Sa, El, Pr	Discussed with team; no direct action taken, more so approach in next year	Ex, Ju, Co	None	n/a	ANLNU	CG	O1, O2, O3
OHCHR 15	O1, 2	y	Finding a balance between NGO and UN compatible messages	N-U	Recomm. reviewed, considered and thought about	Sa, El, Pr	Discussed with team, agreed that action not possible to take	Ex, Ju, Co	None	n/a	ANLNU	CG, OC	O1, O2
OHCHR 16	O1, 2	y	Adopt more specific messages on government pledges	N-U	Recomm. reviewed, considered and thought about	Sa, El	Discussed with team, agreed that action not possible to take	Ex, Ju, Co	None	n/a	ANLNU	OC	O1, O2

Legend

Titles: So: Source, An: Anticipated (Y=yes, N=No), Att: attribute, How: how use occurred, Inf: Influence, Ver: Verification

Source: I1, 2, 3, 4= ICRC interviewees, O1,2,3= OHCHR interviewees

Type: In: instrumental, Pr: process, Con: concept, N-U: non-use, Sy: symbolic

Individual attributes: At: attitude, Sa: salience, E: elaboration, Pr: priming, Ka: knowledge acquisition, Sa: skills acquisition

Interpersonal attributes: Ex: exchange, Co: consensus, Ju: justification, Pe: persuasion:

Collective attributes: PRC: practice change, POC: policy change

How use occurred: ALU: anticipated linear use, ALNU: anticipated linear non-use, ULU: unanticipated linear use, ULNU: unanticipated linear non-use, ANLU: anticipated non-linear use, ANLNU: anticipated non-linear non-use, UNLU: Unanticipated non-linear use, UNLNU: Unanticipated non-linear non-use

Influences: C: context, CG: communication goals and ambitions, E: evaluation policies and institutions, F: funding, OC: organisational context, FI: field, P: people

Verification: D: documentation, I1, 2, 3, 4= ICRC interviewees, O1,2,3= OHCHR interviewees

6.7. References

Albaek, E. (1995). Between knowledge and power: Utilization of social science in public policy making. *Policy Sciences*, 28(1), 79-100.

Alkin, M. C., Daillak, R., & White, P. (1979). *Using evaluations: Does evaluation make a difference.* Beverly Hills, CA: Sage Publications.

Alkin, M. C., & Taut, S. M. (2002). Unbundling evaluation use. *Studies in Educational Evaluation*, 29(1), 1-12.

Ayers, T. D. (1987). Stakeholders as partners in evaluation: A stakeholder-collaborative approach. *Evaluation and Program Planning*, 10, 263-271.

Bamberger, M., Rugh, J., & Mabry, L. (2011). *RealWorld evaluation: Working under budget, time, data, and political constraints*. Newbury Park, CA: Sage Publications.

Bhatt, G. D. (2001). Knowledge management in organizations: examining the interaction between technologies, techniques, and people. *Journal of knowledge management*, 5(1), 68-75.

Birkland, T. A. (2014). *An introduction to the policy process: Theories, concepts and models of public policy making*. New York, NY: Routledge.

Braverman, M. T., & Arnold, M. E. (2008). An evaluator's balancing act: Making decisions about methodological rigor. In M. T. Braverman, M. Engle, M. E. Arnold & R. A. Rennekamp (Eds.), *Program evaluation in a complex organizational system: Lessons from Cooperative Extension*. *New Directions for Evaluation*, 120, 71–86.

Broom, G. M., & Dozier, D. M. (1990). *Using research in public relations: Applications to program management*. Englewood Cliffs, NJ: Prentice-Hall.

Caracelli, V. J. (2000). Evaluation use at the threshold of the twenty‑first century. *New directions for evaluation*, 2000(88), 99-111.

Christie, C. A. (2007). Reported influence of evaluation data on decision makers' actions: An empirical examination. *American Journal of Evaluation*, 28, 8-25.

Christie, C., & Fierro, L. (2012). Evaluation policy to implementation: An evaluation of scientifically based research in practice. *Studies in Educational Evaluation*, 38: 65–72.

Ciarlo, J. A. (Ed.). (1981*). Utilizing evaluation: Concepts and measurement techniques.* Beverly Hills, CA: Sage Publications.

Clay, E. J., & Schaffer, B. B (eds.) (1986). *Room for Manoeuvre, An Explanation of Public Policy in Agriculture and Rural Development*. London: Heinemann.

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: a new perspective on learning and innovation. *Administrative science quarterly*, 128-152.

Contandriopoulos, D., & Brousselle, A. (2012). Evaluation models and evaluation use. *Evaluation*, 18(1), 61-77.

Cousins, J. B. (1995). Assessing program needs using participatory evaluation: A comparison of high and marginal success cases. In J. B. Cousins, & L. M. Earl (Eds.), *Participatory evaluation in education: Studies in evaluation use and organizational learning* (pp. 55-71). London: Falmer.

Cousins, J. B., Goh, S., Clark, S., & Lee, L. (2004). Integrating evaluative inquiry into the organizational culture: A review and synthesis of the knowledge base. *Canadian Journal of Program Evaluation*, 19(2), 99-141.

Cousins, J. B., Goh, S. C., Elliott, C. J., & Bourgeois, I. (2014). Framing the capacity to do and use evaluation. In J.B. Cousins, & I. Bourgeois (Eds.), *Organizational capacity to do and use evaluation. New Directions for Evaluation*, 141, 7–23.

Cousins, J. B., & Leithwood, K. A. (1986). Current empirical research in evaluation utilization. *Review of Educational Research*, (536), 331-364.

de Certeau, M. (1990). *L'invention du quotidien. 1. arts de faire* (1ère édition 1980). Paris: Gallimard.

Dozier, D. M., Grunig, L. A., & Grunig, J. E. (2001). Public relations as communication campaign. In R.E. Rice & C.K. Atkin, (Eds.), *Public communication campaigns*. 3rd Edition (pp. 231-248). Newbury Park, CA: Sage Publications.

Dubin, R. (1976). Theory building in applied areas. In M.D. Dunnette (Ed.), *Handbook of Industrial/organizational psychology* (pp. 17-39). New York, NY: Rand McNally.

Dwyer, S. C., & Buckle, J.L. (2009). The Space Between: On Being an Insider-Outsider in Qualitative Research. *International Journal of Qualitative Methods* 8:54–63.

Fitzpatrick, J. L. (2012). An introduction to context and its role in evaluation practice. *New Directions for Evaluation*, *2012*(135), 7-24.

Gillespie, A., & Cornish, F. (2014). Sensitizing questions: A method to facilitate analyzing the meaning of an utterance. *Integrative Psychological and Behavioral Science*, 48(4), 435-452.

Greene, J. G. (1988). Stakeholder participation and utilization in program evaluation. *Evaluation Review*, (12), 912-116.

Gregory, A., & White, J. (2008). Introducing the Chartered Institute of Public Relations' initiative. In B. van Ruler, A. Tkalac-Verčič & D. Verčič (Eds.), *Public relations metrics* (pp. 307–17). New York, NY: Routledge.

Hall, M. (2008). The effect of comprehensive performance measurement systems on role clarity, psychological empowerment and managerial performance. *Accounting, Organizations and Society,* 33 (2-3), 141-163.

Hall, M. (2011). Do comprehensive performance measurement systems help or hinder managers' mental model development? *Management Accounting Research*, 22 (2), 68-83.

Henry, G. T., & Mark, M. M. (2003). Beyond use: Understanding evaluation's influence on attitudes and actions. *American Journal of Evaluation*, 24, 293–314.

Højlund, S. (2014a). Evaluation use in the organizational context–changing focus to improve theory. *Evaluation*, 20(1), 26-43.

Højlund, S. (2014b). Evaluation use in evaluation systems–the case of the European Commission. *Evaluation*, 20(4), 428-446.

Johnson, K., Greenseid, L. O., Toal, S. A., King, J. A., Lawrenz, F., & Volkov, B. (2009). Research on Evaluation Use A Review of the Empirical Literature From 1986 to 2005. *American Journal of Evaluation*, 30(3), 377-410.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness* (No. 6). Chicago, IL: Harvard University Press.

Julnes, P. D. L., & Holzer, M. (2001). Promoting the utilization of performance measures in public organizations: An empirical study of factors affecting adoption and implementation. *Public Administration Review*, 61(6), 693-708.

Juma, C., & Clark, N. (1995). Policy research in sub-Saharan Africa: An Exploration. *Public Administration and Development*, Vol.15, 121–137.

Katz, D., & Kahn, R. L. (1966). *The social psychology of organizations*. New York, NY: Wiley.

Kerstetter, K. (2012). Insider, outsider, or somewhere in between: The impact of researchers' identities on the community-based research process. *Journal of Rural Social Sciences*, 27(2), 99–117.

Kirkhart, K. (2000). Reconceptualizing evaluation use: An integrated theory of influence. *New directions for evaluation 2000*, no. 88 (2000): 5-23.

Lennie, J., & Tacchi, J. (2013). *Evaluating Communication for Development: A Framework for Social Change.* New York, NY: Routledge.

Leviton, L. C., & Hughes, E. F. (1981). Research On the Utilization of Evaluations A Review and Synthesis. *Evaluation review*, 5(4), 525-548.

Levy, C. (1992). Gender and the environment: the challenge of cross-cutting issues in development policy and planning. *Environment and Urbanization*. 4.1: 134-149.

Luhmann, N. (1990). *Essays on self-reference*. New York, NY: Columbia University Press.

Macnamara, J. (2006). Two-tier evaluation can help corporate communicators gain management support. *PRism*, *4*(2). Retrieved from: http://www.prismjournal.org/fileadmin/Praxis/Files/Journal_Files/Evaluation_Issue/COMMENTARY_MACNAMARA.pdf

Manheim, J. (2011). *Strategy in information and influence campaigns: how policy advocates, social movements, insurgent groups, corporations, governments, and others get what they want*. New York, NY: Routledge.

Mannell, J. (2014). Adopting, manipulating, transforming: Tactics used by gender practitioners in South African NGOs to translate international gender policies into local practice. *Health & place*, *30*, 4-12.

Mark, M. M., Cooksy, L. J., & Trochim, W. M. (2009). Evaluation policy: An introduction and overview. *New Directions for Evaluation*, 2009(123), 3-11.

Mark, M. M., & Henry, G. T. (2004). The mechanisms and outcomes of evaluation influence. *Evaluation*, 10(1), 35-57.

Mason, J. (2002). *Researching your own practice: The discipline of noticing*. London: Routledge.

Mathison, S. (1994). Rethinking the evaluator role: Partnerships between organizations and evaluators. *Evaluation and Program Planning*, 17(3), 299-304.

Meredith, J. (1992). Theory Building through Conceptual Methods. *International Journal of Operations and Production Management,* 13 (5): 3-11.

Mintzberg, H., & Waters, J. A. (1985). Of strategies, deliberate and emergent. *Strategic Management Journal*, 6, 257–272.

Mohammed, S., & Ringseis, E. (2001). Cognitive diversity and consensus in group decision making: The role of inputs, processes, and outcomes. *Organizational Behavior and Human Decision Processes,* 85(2), 310-335.

Mowbray, C. T. (1992). The role of evaluation in restructuring of the public mental health system. *Evaluation and Program Planning*, 15, 403-415.

O'Neil, G. (2013). Evaluation of international and non-governmental organizations' communication activities: A 15 year systematic review. *Public Relations Review*, 39(5), 572-574.

O'Neil, G. (2015). Implementing communication evaluation methodologies for two international organisations. *Submitted for publication.*

Patton, M. Q. (2000). Utilization-Focused Evaluation. In D. Stufflebeam, G. Madaus & T. Kellaghan (Eds.), *Evaluation Models: Viewpoints on Educational and Human Services Evaluation* (pp. 425-538). Norwell, MA: Springer.

Patton, M. Q. (2008). *Utilization-focused Evaluation.* London: Sage Publications.

Patton, M. Q., Grimes, P. S., Guthrie, K. M., Brennan, N. J., French, B. D., & Blyth, D. A. (1977). In search of impact: An analysis of the utilization of federal health evaluation research. In C. H. Weiss (Ed.), *Using social research in public policy making* (pp. 141-163). New York, NY: D. C. Heath and Company.

Radnor, H. (2001). *Researching your professional practice: Doing interpretive research*. Buckingham: Open University Press.

Russ-Eft, D., Atwood, R., & Egherman, T. (2002). Use and non-use of evaluation results: Case study of environmental influences in the private sector. *American Journal of Evaluation*, 23(1), 19-31.

Saaty, T. L. (1999). *Decision making for leaders: the analytic hierarchy process for decisions in a complex world* (Vol. 2). Pittsburgh, PA: RWS publications.

Scott, R. W. (1995). *Institutions and Organizations*. Newbury Park, CA: Sage Publications.

Scriven, M. (1986). New frontiers of evaluation. *Evaluation Practice*, 7, 7-44.

Shulha, L. M., & Cousins, J. B. (1997). Evaluation use: theory, research, and practice since 1986. *American Journal of Evaluation*, 18(3), 195-208.

Stacks, D., & Michaelson. D. (2010). *A practitioner's guide to public relations research, measurement and evaluation*. New York, NY: Business Expert Press.

Starling, S. (2010). Monitoring and evaluating advocacy: lessons from Oxfam GB's Climate Change campaign, *Development in Practice*, *20* (2), 277-286.

Sutton, R. (1999). *The policy process: an overview.* London: Overseas Development Institute.

Torres, R. T., Preskill, H. S., & Piontek, M. E. (1996). *Evaluation strategies for communicating and reporting: Enhancing learning in organizations*. Thousand Oaks, CA: Sage.

Trochim, W.M.K. (2009). Evaluation policy and evaluation practice. In W.M.K. Trochim, M.M. Mark & L.J. Cooksy (Eds.), *Evaluation policy and evaluation practice*. *New Directions for Evaluation, 123,* 13–32.

Vandenbosch, B., & Higgins, C.A., (1995). Executive support systems and learning: a model and empirical test. *Journal of Managements Information Systems,*12, 99-130.

Vandenbosch, B., & Higgins, C.A. (1996). Information acquisition and mental models: an investigation into the relation between behaviour and learning. *Information Systems research* 7:198-214.

Watson, T. (1997). Measuring the success rate: Evaluating the PR process and PR programmes. In P. J. Kitchen (Ed.), *Public relations: Principles and practices* (pp. 283–99). Boston, IL: International Thomson Business Press.

Watson, T., & Noble, P. (2007). *Evaluating public relations*. London: Kogan Page.

Weiss, C.H. (1979). The many meanings of research utilization. *Public Administration Review*, 39, 426-431.

Weiss, C.H. (1981). Measuring the use of evaluation. In Ciarlo, J. A. (Ed.). (1981*). Utilizing evaluation: Concepts and measurement techniques* (pp. 17-31). Beverly Hills, CA: Sage Publications.

Weiss, C.H. (1998). *Evaluation: Methods for Studying Programs and Policies* (2nd ed.). Newbury Park, CA: Sage Publications.

Weiss, C.H., Murphy-Graham, E., & Birkeland, S. (2005). An alternate route to policy influence how evaluations affect DARE. *American Journal of Evaluation*, 26(1), 12-30.

Xavier, R., Patel, A., Johnston, K., Watson, T., & Simmons, P. (2005). Using evaluation techniques and performance claims to demonstrate public relations impact: An Australian perspective. *Public Relations Review, 31*(3), 417–24.

Zerfass, A. (2008). A framework for managing and evaluating communication strategies: The Corporate Communications Scorecard. In van Ruler, B., Tkalac-Verčič, A., & Verčič, D. (Eds.). *Public relations metrics* (pp. 139-153). New York, NY: Routledge.

Interlude 4

Article 3 provided an in-depth study on the use of evaluation findings of the two communication campaigns whose evaluation was the focus of the previous article 2. Article 3 identified instances of use and non-use in these organisations which were mainly unanticipated and instrumental in nature. Article 3 provided further evidence to support articles 1 and 2 on the factors that influence the evaluation process, and in this case, the use of evaluation findings. As for article 2, the strongest influences were found to be internal, organisational context and communication goals and ambitions, which could enable or impede use, depending upon the situation.

Article 3 confirmed the standpoint taken at the start of this thesis that the evaluation process cannot be viewed in isolation; it was an interrelated part of the organisation and any change it provoked had to be considered in relation to other influences drawn from the implicated people, organisational settings and contexts.

In the Conclusion chapter of this thesis, the evidence produced by the three articles of this thesis are summarised and the three research questions responded to. Theoretical implications, methodological lessons, recommendations for the practice of communication evaluation and future research, and broader implications are provided.

7. Conclusion

Before the main findings and their implications are discussed, it is helpful to recall the specifications of this PhD thesis. The central question of this thesis was to assess the extent to which communication evaluation is possible within IOs. This central question was developed into three specific questions, in brief: 1) the appropriateness and feasibility of communication evaluation methodology for IOs; 2) the influence of internal and external factors in its implementation and; 3) the use of communication evaluation findings in IOs. These questions were responded to in three distinct but interlinked empirical studies that form the core of this thesis as articles 1, 2 and 3. A conceptual framework for communication evaluation of IOs created by this author has guided and framed this thesis. The structure of this thesis can be visualised in section 1.3, Figure II. The conceptual framework can be visualised in section 3.1, Figure I.

This conclusion chapter is organised as follows. Section 7.1 reviews how the evidence presented in the articles respond to the three research questions. Section 7.2 considers the theoretical implications of the findings. Section 7.3 draws some conclusions on methodologies used in this thesis. Section 7.4 provides recommendations for the practice of communication evaluation within IOs. Section 7.5 suggests future directions for further research in the subjects covered by this thesis. Section 7.6 provides a reflection on the broader implications of the findings of this thesis.

7.1. Summary of research questions and findings

The central question of this thesis asked to what extent communication evaluation in IOs is possible. Of course, even before reading this thesis, a reader could surmise "it is possible". However, where this thesis aimed to provide a contribution was in understanding *to what extent* is it possible. As the responses below to the specific questions illustrate, the possibility to evaluate communication activities to *the extent* that it is of value and use to organisations depends upon the evaluation approach adopted and

the influences encountered. The suggested approach and strategy to counter factors of influence is discussed below.

7.1.1. Specific question 1: What evaluation principles, methods and procedures are appropriate and feasible for IOs to evaluate their global communication programmes and campaigns?

This thesis found that the methodology appropriate for IOs to evaluate their global communication programmes and campaigns is a combination of a robust design and diverse methods matched to the outcomes being evaluated, implemented with a pragmatic, adaptive and participatory approach.

The support of this thesis for a pragmatic and adaptive approach implies favouring an epistemological orientation for communication evaluation within IOs, resembling closest to an eclectic mixed-methods paradigm. As noted at the start of this thesis, epistemological discussions have not been addressed substantially in the communication evaluation literature to date (section 2.5.6). Although in this study, this is an orientation that emerged in article 2 rather than being dictated, it does not rule out that other epistemological orientations could be appropriate, for example, a positivist-quantitative orientation for campaign evaluation in the health field.

However, as this study showed, the range, complexities and ambitions of communications of IOs would indicate that any epistemological position that insists on a given method, design or approach, would not be appropriate. As Hall (2014) suggested, methodologies can be deemed inappropriate for valid methodological reasons (e.g. where the method is not matched to the data to be collected) but should not be ruled out on ideological grounds (e.g. a given method is deemed superior over others). In this respect, this study shows that a universal approach for evaluating communication activities of IOs that sets out standard methods and measures would not be feasible. More so, it was found that a process of conceptualisation would be appropriate that matches the given communication activities, organisational settings and context to an appropriate

methodology. This is expanded upon further below in section 7.4 when considering recommendations for future communication evaluation practice.

The three key methodological elements mentioned in specific question 1 are now discussed.

Principles: The broad principles for communication evaluation methodology, as used in article 1 were shown to be valid for IOs, even if they could be considered as not specific for IOs, but more so minimum expectations for evaluation-related actions for communication activities of all sectors. Beyond these overarching principles, this study and the literature indicates that IOs will increasingly adopt more formal evaluation policies that could also include specific evaluation principles of a given organisation, which is the case of such policies to date, for example for the ILO evaluation policy (ILO, 2012). In this regard, in line with the above argument, any such principles could be incompatible with communication evaluation methodology if they advocate a preference for a given epistemological position and consequently prescribe methods to be used.

Methods: The lack of diversity in methods for communication evaluation was confirmed by this thesis (article 1). It was found that the broad range of communication activities carried out by IOs implies that equally a broad range of evaluation methods is required. It was found that a standard palette of methods could not be recommended; more so that each communication output/outcome would need to be considered individually (or in relevant groupings) and matched to appropriate method(s). In evaluating communication campaigns of IOs, this thesis used a nested analysis mixed methods strategy that proved appropriate for the range of activities being evaluated (article 2). Challenges in the use of methods were more related to issues of feasibility, that is, the ability of organisations (or their evaluators) to deploy methods to match the complexity of their activities, organisational settings and contexts, as discussed for the next specific question.

Procedures: Evaluation within IOs was found to be dominated by logical and linear processes that are reflected in their procedures and have been reinforced by results-based

management systems in their operationalisation. This thesis found that for certain aspects, a logical and linear process facilitated evaluation and its analysis, for example, the categorisation of communication effects from input to impact (section 2.5.3., Table I); and the procedural steps of the evaluation process adopted by the conceptual framework of this thesis (section 3.1., Figure I). Where this thesis found it would be appropriate to deviate from such an approach was for implementation procedures, notably in the need to use pragmatic, adaptive and participatory approaches (article 2). It was found that an evaluation and its evaluator(s) has to adopt such an approach in order to respond to the ambitions of the communication activities, the challenges faced by the factors of influence and to build confidence of staff in the evaluation methods and eventual findings. As reported by previous research, this was essentially the "evaluator's balancing act" (Braverman & Arnold, 2008, p. 71). Evidence from article 3 also supported the notion that adopting a participatory approach does facilitate the use of evaluation findings.

In addition to these three methodological elements, a fourth deserves to be mentioned, that is, the evaluation design. It was found that the choice of the evaluation design proved to be crucial in the ability to adhere to the above-mentioned evaluation principles, to use more appropriate methods and to provide the possibility of the evaluation to have an input into the planned communication strategies and activities (articles 1 and 2). This thesis found that the pre-post design offered these advantages over the dominant post-only design. However, articles 1 and 2 found that the possibility to use a pre-post design was reliant on evaluation being considered and integrated within the communication function early in the communication planning process, thus indicating the influence on this and possibly other factors, as discussed in the next section. This thesis found no evidence to support the widespread use of experimental and quasi-experimental designs (pre-post with control groups) for communication evaluation of IOs. This is not to deny that the use of such designs is appropriate for some communication evaluation, for example, in the evaluation of health campaigns with precise behaviour change objectives, where it is the prevailing design (Valente, 2001). However, given the range and complexity of communication activities of IOs identified, such a design would neither be appropriate nor feasible for the majority of evaluations.

The above findings of this thesis focus on the appropriateness of evaluation methodology for communication evaluation of IOs. However, feasibility, that is, what in practice is possible to do, also has to be considered. For example, the adoption of a given method could be appropriate, but it may not be feasible to implement. Similarly, an evaluation design may be feasible, but not appropriate. This thesis found that issues of feasibility were often influenced by factors outside of the evaluation process, predominantly internal as discussed further in the next section.

The findings on this question added to the existing literature, notably as the first known empirical study on communication evaluation within IOs and the insights provided on the appropriateness and feasibility of methodologies and their implementation. The process of conceptualisation cited has been emphasised previously by Grunig (2008). The findings on the dominant nature of the logical and linear processes have been previously confirmed (Mayne, 2007; Rogers, 2008; Sanderson, 2000). The lack of diversity in methods found by this thesis has been previously noted in the literature (Coffman, 2002; Grunig, 2008; Jelen, 2008), as also have the advantages of the pre-post evaluation design (Broom & Dozier, 1990; Lehmann, 2007). Further, the findings provide a broader perspective on methodology for the communication evaluation literature that has been lacking to date.

7.1.2. Specific question 2: To what extent do the broader contextual, organisational and human factors influence the ability of IOs to evaluate their global communication programmes and campaigns?

The ability of IOs to evaluate their global communication programmes and campaigns was found to be influenced by broader contextual, organisational and human factors. All three articles supported the notion that mainly internal factors influenced the ability of organisations to evaluate, which could be both enabling or impeding, depending upon the situation. These factors were found not only to influence the ability to evaluate, but the broader evaluation process, from conceptualisation of methodology to use of evaluation findings. Organisational factors that were found to be more important in their influence

than others were the organisational context and communication goals and ambitions. Table I details the factors of influence of the conceptual framework of this thesis and provides examples of their influence, both enabling and impeding, drawn from the three empirical studies.

Table I: Summary of factors influencing the evaluation process		
	Enabling / impeding	*Examples of influence found*
Internal factors:		
Evaluation policies and institutions	Enabling	• Policies regularised that encourage take up of evaluation • Evaluation institution and policies promulgate standard terminology, process and focus on appropriate level of effect (outcome) • Evaluation institution and policies encourage evaluation practices in functions such as communications
Organisational context	Impeding	• Network structure of organisations relying on partners for aspects of the evaluation and access to their audiences • Organisational priorities limit range of communication messages
	Enabling	• Emphasis placed on evaluation in organisations and appropriate level of effect (outcome) • Flexibility of communication managers to launch evaluations
Communication goals and ambitions	Impeding	• Lack of integration of evaluation in communication function • Complexity of global communication and campaign models with multiple objectives and audiences
	Enabling	• Strategies provide guidance for what to be evaluated
External factors:		
Context	Impeding	• Inability to access all targeted audiences • Complex and varied environments • Rapidly evolving communication landscape
Field	Enabling	• Promote diversity of evaluation methods • Promote standard terminology, process and focus on appropriate level of effect (outcome) • Provide comparable organisations
	Impeding	• Limited guidance on evaluating communication contribution to organisation & societal levels • Limited guidance on evaluating emerging communication activities, e.g. online activities
Internal/External factors:		
Funding	Impeding	• Availability of budgets for evaluation and more diverse methods
People	Impeding	• Inability of external evaluation consultants to use diverse methods • Evaluation not part of regular tasks of communication staff
	Enabling	• Communication staff engaged in evaluation • Communication staff using evaluation findings

This thesis found that the majority of factors could be both enabling and impeding influences on the evaluation process, as illustrated in Table I. A factor could alternate between impeding and enabling depending upon the situation. For example, organisational context could be impeding in that their network-based structures for communicating meant that partners were relied upon for certain aspects of the evaluation, such as some data collection, which could not be fully controlled. At the same time, the organisational context could be enabling by the emphasis placed by the organisation on evaluation for functions such as communication.

Factors were found to vary in their strength and subtlety of influence. For example, evaluation policies and institutions were found to be strong in organisations where it was institutionalised and weaker in organisations where it was not (article 1 and 3). Organisational context could play a decisive role in determining the communication and consequent evaluation approaches used, through its structures and overall strategies (article 2). At the same time, influence of a factor could vary in its subtlety. For example, for the field factor, its influence could be subtle and not obvious through the consistent and constant promotion of standard terminology and processes by the evaluation field; at the same time, it could be more explicit when organisations directly drew from the field, such as in comparisons to like-minded organisations (articles 2 and 3).

Factors of influence were found to be interlinked. For example, the evidence indicates that the external context influenced the organisational context in the setting of priorities and strategies, which the communication function then responded to in their selection of objectives and target audiences. People, communication staff or external evaluation consultants, implemented procedures and methods drawn from or in conformity with other factors, such as the evaluation field and the organisation's evaluation policies and institutions. Communication ambitions were limited by the funding available. These interlinkages are described further in section 7.2 below.

The findings for this question expanded upon the current literature which has tended to focus on a single factor, such as context (Fitzpatrick, 2012) or the influence of factors on

a given component of the evaluation process, such as use (Alkin & Taut, 2002; Mark & Henry, 2004). The three articles of this thesis show support in favour of considering these factors as interlinked and varying in their strength and subtlety of influence, consistent with a systems perspective of organisations and their parts (Banathy, 2000; Bertalanffy, 1969).

7.1.3. Specific question 3: How are communication evaluation findings used within IOs and what factors enable and impede their use?

Within the two IOs studied, instances of both use and non-use of communication evaluation findings were identified in article 3. Instrumental use, that is, direct use of findings, was found to be dominant followed by non-use, that is, where a recommendation of findings has been considered and not taken up. Instances of conceptual use, that is, where understanding has been affected but no direct action taken, followed. Instances of use decreased when moving away from the communication managers who had commissioned the evaluations, both in time and in distance. Participation of staff in an evaluation was found to be linked to their use of findings.

Those evaluation findings where communication staff acted upon were found to be mainly concerned with improving the efficiency and relevance of future communication activities. This interest of staff in issues of efficiency and relevance was already seen in the methodology and implementation components of evaluation process (article 2).

Use of findings was found to be unpredictable with identified instances being mainly unanticipated, that is, not a stated recommendation but drawn explicitly or implicitly from the findings. It was found that communication staff assigned different meaning to the evaluation findings, adjusting and interpreting findings in opportunistic and unexpected ways. Article 3 also indicated that the meaning assigned to an evaluation and its findings could vary from the routine to the symbolic, and thus impact on its use; for the former being more mixed and the latter more distinct.

In the two IOs studied, evaluation use was found to travel from the individual to the organisational level in a predominantly non-linear fashion, interconnected, overlapping and bringing about change both in a formal and informal manner but never in a vacuum devoid of other influences. Use was found not always to happen immediately and directly. Each instance of use could be categorised into a type of use (a process categorisation of eight types), illustrating that use in these two cases was predominantly unanticipated and could be equally linear or non-linear, as follows (ranked and using travel analogies):

   -*Unexpected hop*: Use was not anticipated and occurred in a linear way.

   -*Unforeseen foray:* Use was not anticipated and occurred in a non-linear way (equal first).

   -*Direct route*: Use was anticipated and occurred in a linear way.

   -*A planned ramble*: Use was anticipated and occurred in a non-linear way (equal second).

   -*Expedition starts/stops*: Use was anticipated, did not occur and happened in a non-linear way.

   -*Travel plans cancelled*: Use was anticipated, did not occur and happened in a linear way.

   -*Surprise trip deferred*: Use was not anticipated, did not occur and happened in a non-linear way.

   -*Unannounced stop-over skipped*: Use was not anticipated, did not occur and happened in a linear way.

Over time, it was also possible that certain aspects of the findings would become consumed in communication policies and practices formed from multiple sources, with their origin lost to most staff.

Factors that enabled and impeded the use of findings were found to be the same as those that influenced other components of the evaluation process, namely internal factors of organisational context and communication goals and ambitions, as seen above in Table I. As for the other components, these factors both enabled or impeded use, depending upon

the situation. All instances of use originated in individuals (staff) and their role was crucial in placing the giving evaluation finding in the appropriate discussion or process and advancing it further. At the same time, staff tended to use findings that strengthened their beliefs and complemented their existing information.

The findings for this question expanded upon and in some cases diverged from the existing literature. Findings on the dominance of instrumental and non-use, the phenomenon of unanticipated use, the non-linear nature of change have not been widely reported in empirical studies of evaluation use (Ciarlo, 1981; Cousins & Leithwood, 1986; Johnson et al, 2009). This study was the first known successful operationalisation of the influence framework of Henry and Mark (2003 & 2004), as was the development of process-categorisation with eight distinct categories (table III, article 3). The emphasis on factors of influence extended previous studies that have tended to focus on the evaluation process and products as influences of use with only more recently context and organisational settings given more attention (Contandriopoulos & Brousselle, 2012; Højlund, 2014).

7.2. Theoretical implications

This thesis developed a new conceptual framework for communication evaluation specific for IOs and is presented again as a reminder below (Figure I). The three empirical studies (articles 1, 2, 3) was an opportunity to test this framework. It was found that the framework was an accurate representation of the key components and factors of influence although it did not show the nuances, complexities and interrelations between its different parts, as expanded upon below.
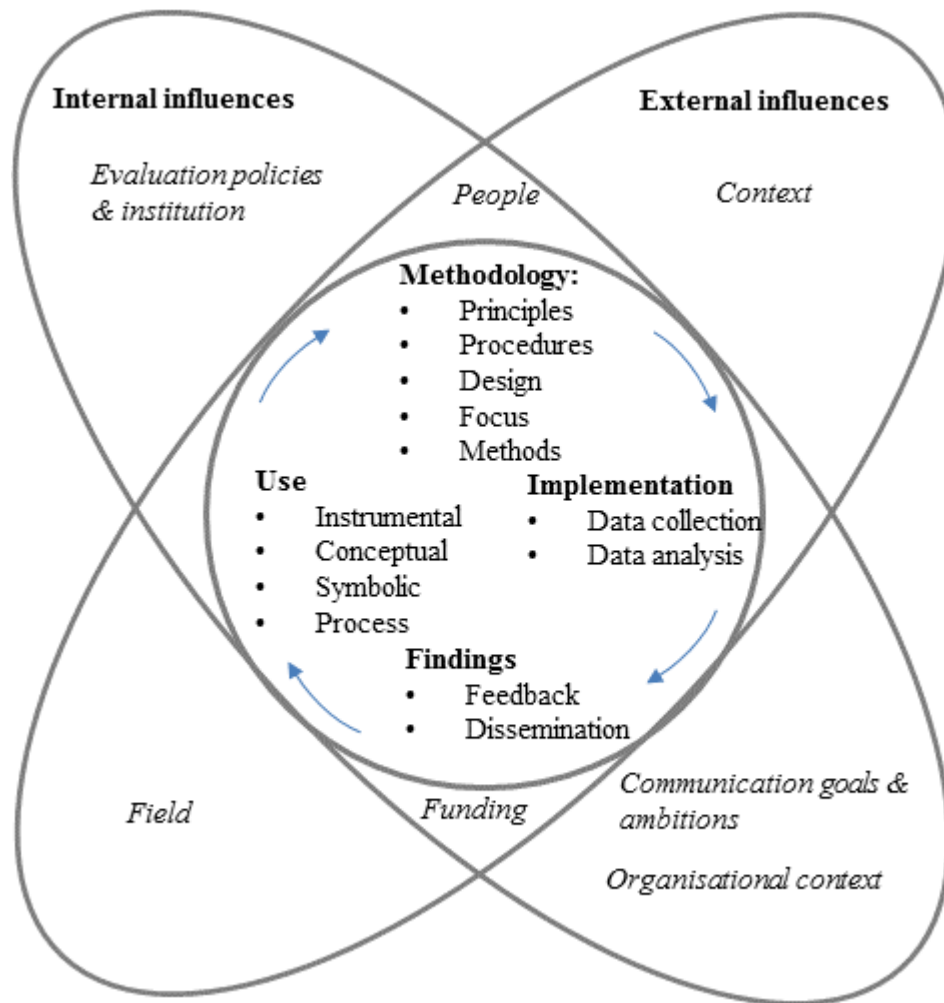
Figure I: MIFU conceptual framework for communication evaluation for IOs (reminder)

In this regard, one key aspect was the weighting of the factors of influence. The three articles indicated that the two factors of communication goals and ambitions and organisational context were the two most important influences. At the same time, factors were found to be enabling or impeding or a combination of both, as illustrated in the table below (a plus sign indicates enabling and a minus sign impeding; with size indicating extent of influence – less or more significant). Some factors were found to be consistent in their influence across the four components, such as evaluation policies and institutions that was an enabling influence throughout the evaluation process. Context was impeding throughout but strongest in the implementation component. People and field were both enabling and impeding in the first two components (methodology and implementation)

and were then only enabling for the last two components (findings and use). There were no additional factors of influence identified with the caveat that the communication goals and ambitions factor was found to be larger in scope than originally conceived, that is, it was found to include elements related to the communication function that creates the goals and ambitions.

Table II: Positive/negative influence of factors on components of evaluation process				
Components / Factors	Methodology	Implementation	Findings	Use
Communication goals and ambitions	+-	+-	+-	+-
Organisational context	+-	+-	+-	+-
Evaluation policies and institutions	+	+	+	+
Context	-	-	-	-
Field	+-	+-	+	+
Funding	-	-	-	-
People	+-	+-	+	+

The four components and their cyclical nature were confirmed by this thesis. As stated above, categorising the evaluation process into four components facilitated the research and analysis for this thesis and broadly represented the reality as to how the evaluation process occurred in the two organisations studied. The thesis also showed that linkages existed between these four components that were not illustrated in the framework. For example, use in the form of process use could occur throughout the four components; feedback, an element of the findings components also occurred often in the implementation component.

Further, it was found that the exchange between the factors of influence and components of the evaluation process was not only one-way from the factors to the components but also the contrary. A dynamic, interconnected and overlapping flow of influence was seen. For example, people, such as staff, were an influence on all aspects of the evaluation process but they were also influenced positively by their participation in the evaluation (e.g. by acquiring new knowledge and skills). Equally, the factor of communication goals and ambitions was an important influence on the evaluation process but then in return, select evaluation findings influenced future goals and ambitions.

The thesis also showed that linkages existed between the factors of influence, with factors feeding to and from each other. Figure II illustrates the main linkages between the factors as found by this thesis. The mainly external factors of context, funding and field influenced internal factors such as the organisational context and communication goals and ambitions that consequently influenced people and evaluation policies and institution, sometimes directly or indirectly.
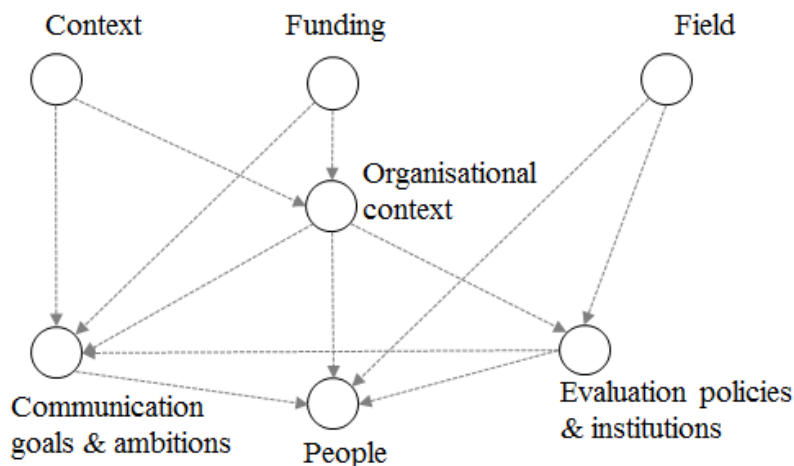
Figure II: Main linkages between factors of influence

The findings of the thesis support the relevance of systems theory in understanding communication evaluation in IOs. The evidence supports the original standpoint taken by this author (section 3.1) that organisations are made up of interrelated parts, adapting and adjusting to the environment in which they operate. Communication evaluation cannot be viewed in isolation, more so it is interlinked to other processes and parts of the

organisation. The communication evaluation process was shown to be an example of how the organisation interacts both internally and externally, and adjusts itself based on the inputs provided by evaluation. The findings of this thesis suggest that organisations have to be aware of the dynamic nature of their parts concerning evaluation and be able to react and adjust accordingly, in order to draw value from the evaluation process and the findings it produces. Systems theory provided the main theoretical frame of reference for the thesis, notably in supporting the understanding of interrelationships, multiple perspectives and boundaries within the communication evaluation process and between it and the organisations concerned.

7.3. Methodological lessons

This thesis utilised a variety of methodological approaches and a reflection is provided on the main approaches used.

Article 1 was guided by a systematic review methodology. This method had the advantage that it provided a comparison of current evaluation practices across IOs for a given period of time and provided a broad understanding of the trends and patterns in these practices. The corpus studied was evaluation reports of different institutions and consequently content and quality differed from the normal corpus studied by this method, which is scientific evidence (from journal articles or other sources). In this regard, the extraction of comparable data was aided by having a detailed coding protocol and double-coding as a check on reliability.

The limitations seen in using this method was that it provided only a partial view of evaluation activities within IOs, as only two outputs or products of the evaluation process were considered, that is, the evaluation report and evaluation policies and guidelines. It could be argued that these two outputs are the most important visible elements of an evaluation process. However, through interacting with organisations in articles 2 and 3, this author understood that not all evaluation activities being carried out were captured by these two outputs, such as monitoring activities and more informal reviews and

assessments. Another limitation seen was that the systematic review produced a picture of these organisations' practices in communication evaluation but could not fully explain the phenomena observed. For example, the predominance of post-only evaluation design was seen yet it was not possible to fully understand why this was so; the corpus provided some insights as select evaluation reports detailed their design choices. In this sense, using the systematic review methodology provided a good starting point for this thesis with its findings complemented by the use of other methods in articles 1 and 2.

The author drew on the "insider-outsider" research approach for Articles 2 and 3. This approach supported the author in the way that the research was carried out and served as an overall anchoring for the relationship between the author and the organisations studied. A complexity added to this role, was that as an evaluator, the author was introducing and carrying out evaluation research with the organisations, and then standing back and reflecting upon the approaches, methods and processes used. Overall, the "insider-outsider" research provided a necessary frame for the author to understand, analyse and report on his dual role as an evaluator and researcher.

Article 3 used qualitative methods to investigate in-depth instances of evaluation use based on a typology created by this author from multiple sources. This effectively developed a methodology for determining individual instances of evaluation use, to analyse them from different perspectives and to validate their plausibility. This proved a useful development for this thesis as the current methodology on evaluation use was found to be inappropriate, either based on quantitative surveying for large populations or qualitative discussions on general impressions of use. With both methods, neither was suited to the in-depth study of instances carried out in article 3. An advantage of this method was the rich detail provided on instances, the people, processes and factors of influence involved. The collection of multiple instances of use also facilitated the creation of a process categorisation on how use occurs, e.g. linear, non-linear, anticipated, unanticipated, etc. A challenge seen was in applying the above typology to the experiences described by communication staff in their use of evaluation findings. As a response, the typology was constantly revised based on the discussions with staff that led

to the addition of factors within the typology's dimensions and to contest some underlying assumptions of the typology, such as the linear nature of decision-making.

7.4. Recommendations for communication evaluation practice

As this thesis was a study of a specific function (communication) of a defined population (IOs) and their interaction with a particular phenomenon (evaluation), it is only natural that recommendations can be drawn for future communication evaluation practice of IOs. These recommendations are for four areas: (1) Structure and functions; (2) Conceptualisation; (3) Methods and procedures; and (4) Managing constraints.

Structure and functions: All three empirical studies of this thesis indicated that the integration of evaluation within the communication function was key to adopting more appropriate and effective evaluation practices. IOs should consider how evaluation is incorporated within their communication functions, and in the various operational policies that guide and orientate communication activities, such as job descriptions, project descriptions, communication concept papers and plans of action. Given resource limitations, it would not be expected that new posts for evaluation are created, but more so that evaluation is incorporated into the activities of existing communication staff.

Conceptualisation: As described above, a pre-determined methodology is not appropriate (nor feasible) for the communication evaluation in IOs. It is recommended that before an evaluation action commences, a process of conceptualisation occurs to consider the most appropriate and feasible evaluation approach, methods and implementation for the given communication activity or set of activities. This would seem self-evident but based on the evidence collected by this thesis it does not always seem to be occurring in a systematic and regular manner.

Methods and procedures: As described throughout this thesis, it would be recommended for IOs to adopt a greater diversity of evaluation methods, and in particular to take advantage of new and emerging methods described in this thesis. Aside from favouring a

type of design (pre-post), this author shies away from recommending particular evaluation methods as this thesis found that methods should be selected as a result of the above process of conceptualisation and that any pre-selection or preference for methods or a set of methods drawn from a given epistemological bias should be approached with caution. In terms of procedures, it would be recommended to adopt a pragmatic, adaptive and participatory approach for evaluation in general. In this regard, it is interesting to note the main lesson drawn from discussions with 12 US-based expert evaluators on their practical experiences of carrying out evaluations:

> A primary lesson that transcends each stage is for evaluators to be aware of the choices they have. One should make those choices consciously, considering the context of the program, its state, and stakeholders' information needs and expectations and, given the evaluator's knowledge and expertise in evaluation, what is appropriate and feasible to accomplish in the evaluation. (Fitzpatrick, Christie & Mark, 2009, p. 387)

What this thesis found was that making "choices consciously" involves adopting a certain approach or orientation. This included: the ability to adapt an evaluation approach to the settings, context and demands of relevant staff; the capacity of evaluators to be flexible and ready for reasonable compromises if needed; and the know-how to find ways of involving and consulting with the relevant staff.

Managing constraints: A key finding of this thesis was that the ability of organisations to implement more appropriate and effective evaluation is largely dependent on factors external to the evaluation process. Therefore, organisations have to consider to what extent they can counter and/or capitalise on these outside factors in order to facilitate communication evaluation at its various steps. Evidently, there are limits to the extent to which certain factors of influence can be countered, for example elements of the external context that influence a given communication activity and the consequent evaluation choices. However, it is proposed that each factor should be considered and assessed

accordingly, based on the findings of this thesis that suggest the relative importance of the identified factors of influence.

To support the implementation of these recommendations by communication professionals, the main points have been mapped out in a simplified influence diagram (Figure III), an approach for graphically representing interrelationships among a variety of factors that can be used as a tool for decision-making and analysis (Diffenbach, 1982).

The diagram depicts four groups of questions that needed to be asked during the evaluation process and follows approximately the four components of the evaluation process; 1)Pre-conditions; 2) Conceptualisation; 3) Approach; and 4) Finding value. Pre-condition questions are as stated, i.e. that these points would normally need to be considered before communication evaluation take place. From there, the notion is that if questions are responded to satisfactorily for each group, the interaction of the elements would produce cumulatively a final "result", that is, the evaluation is of value and used by the organisation.

1. Pre-conditions
- Level of integration of evaluation with communication function?
- Evaluation incorporated in job and project descriptions, plans of actions and concept notes?
- Level of support of evaluation policies and institution?
- Availability of necessary budget for evaluation?

Methodology →

2. Conceptualisation
- Level of goals and ambitions defined?
- Matched against appropriate design and methods?
- Type of organisational model to be used?
- Expectation of staff understood?

Implementation

Evaluation of value & used

Use

4. Finding value
- Access to evaluation findings for relevant staff ensured?
- Evaluation findings can contribute to staff knowledge?
- Evaluation findings support more efficient communications?
- Evaluation findings can input into future goals and ambitions?

Findings

3. Approach
- Level of participation of relevant staff ensured?
- Ability to adapt to context?
- Ability to implement diverse methods?
- Flexibility to change methods and data?
- Confidence built in staff of methods, data and consequent findings?

Figure III: Key questions for communication evaluation for IOs

In practical terms, these four groups of questions could work as a type of checklist for communication managers, staff and evaluation consultants to consult and assess their current and/or planned communication evaluation activities.

Finally, drawn from section 7.3, the methods used for this study are of potential interest for communication practitioners and evaluation practitioners in general. The systematic review methodology would be useful for providing an overview of evaluation practices in a given sector or within an organisation or group of organisations (for example a federation or network). The "insider-outsider" approach would be useful for communication and/or evaluation consultants in studying and consequently documenting their experiences and relationships in working with organisations. The methods and typology developed for evaluation use would be of interest for communication and evaluation practitioners in understanding and analysing evaluation use in an organisation.

7.5. Directions for further research

This thesis was the first known in-depth study of communication evaluation within IOs. Therefore, at the general level, more research in this area would be needed to complement the findings of this thesis. In addition, this author has identified the following specific directions for future research.

This was a study of communication evaluation within a particular population (IOs) using a specific set of methods and a theoretical framework. It would therefore be useful to apply this approach with a different population, such as national NGOs, multinationals or government agencies, to see if similar results are found and if this approach is applicable beyond this population. Practically, this may prove difficult to implement, given the time, resources and access to the given population needed. Another alternative would be to replicate certain aspects of this study with other populations, for example, the influence of internal and external factors on the evaluation process of government agencies. This in itself is an area which is lacking theoretical reflection and empirical studies.

This thesis looked at the evaluation process mainly through the lens of actual evaluations carried out. As indicated above, this is an important output or product of the evaluation process, but not the only one. Another output/product is the monitoring data and information that is often produced in parallel, and is predicted to increase given the rise

of automated systems of collection, such as those used for social media and internet communication (part of the so-called "big data"). It would therefore be interesting to apply the same research questions to monitoring; what is appropriate and feasible methodology for communication monitoring in IOs; what is the influence of internal and external factors in its implementation and how are monitoring findings used.

This thesis focused on the activity, campaign and programme level. It is at this level where communication evaluation has focused upon and consequently this thesis has produced findings relevant for this level. However, as indicated in the empirical studies and literature, further research is required to understand the contribution of the communication function to other levels, such as to the organisational and societal levels. This thesis provides some insights into the interrelations between the communication function and its activities, evaluation and the organisation, however this would need to be explored further.

Beyond communication evaluation, two other broader directions are mentioned.

Concerning the use of evaluation findings, it was found that use is often studied as one singular influence on individuals and their organisations. Therefore, rather than studying this isolated input of evaluation findings on policies, it would also be interesting to study a policy process in its totality to understand the relative influence of inputs, including evaluation findings.

Considering evaluation in its conception, implementation and use was central to this thesis. In considering evaluation as a part and process of an organisation and how it interrelates with other parts and factors, this author realised that other concepts must be in a similar situation to evaluation. That is, concepts that are introduced and operationalised in organisations and come with the support of a specific field, system, institution, procedures and expertise. Gender, corporate social responsibility, environmental impact or more applied concepts such as performance measurement or knowledge management come to mind. A comparative study between these concepts, methodologies used, factors

of influence and how their results or other outputs are used in organisations would also be of interest to furthering research.

7.6. Broader implications

In the introduction of this thesis, its main contribution was described as the intersection of three bodies of knowledge: IOs, communication and evaluation, as reproduced in Figure IV below. This chapter has described the specific findings to the questions focused on this intersection. However, this thesis also produced findings of broader interest and implications that are categorised at the intersections of each body of knowledge, as illustrated in Figure IV and described in the next paragraphs.
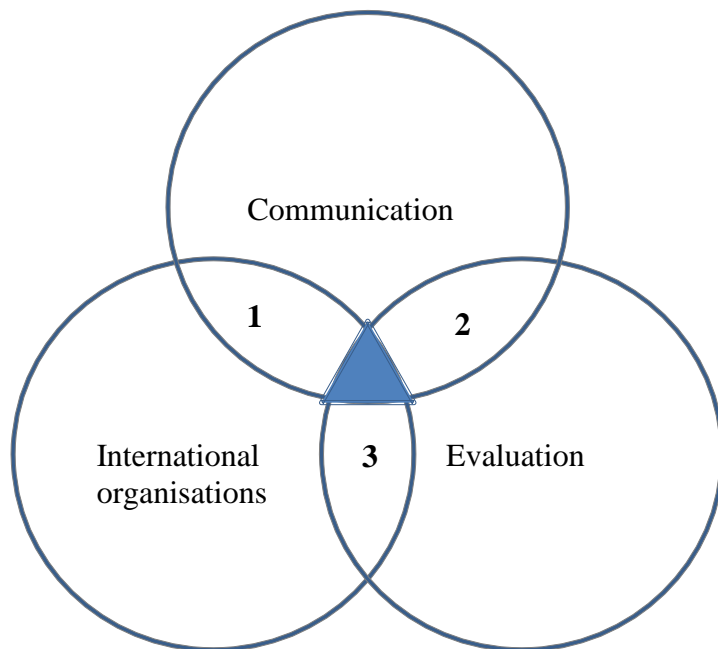
Figure IV: Implication of PhD to intersections of bodies of knowledge

7.6.1. Intersection 1: Communication and IOs

This thesis and the literature confirm that IOs will increasingly use communication as a strategic approach to achieve their goals. Further, it is implied that their ability to influence will increasingly be through "soft" measures such as communication and less through technical assistance, that for many was their original *raison d'être*. At the same

time, communication has found to be increasingly transactional, a dialogue and an exchange where meaning is co-created. Therefore, the extent to which IOs are able to understand and cope with this new reality is still open to question.

7.6.2. Intersection 2: Communication and evaluation

Linked to the above, as communication ambitions of organisations are growing, this thesis found little evidence that the communication evaluation field is yet able to match this change, for example by developing and promoting evaluation methodologies on the contribution of communication to the organisational and societal levels. As organisations are bolder in what they want communication to achieve, evaluation will also have to be able to support this; at the same time, organisations have to be willing to "expose" themselves by putting such ambitions forward for evaluation.

This thesis came to the conclusion that evaluation findings that were of value and used by communication staff were broadly in the area of efficiency and addressed questions such as how can communication activities better reach their audiences and further support their organisations. However, evaluation usually places effectiveness and accountability over these questions, which are essentially centred on efficiency. This author takes into account this point of view but would advocate that this is a legitimate purpose and priority for communication evaluation.

7.6.3. Intersection 3: Evaluation and IOs

This thesis concluded that the most appropriate approach for communication evaluation within IOs is one that is pragmatic, adaptive and participative. In the two evaluations carried out by this author (article 2), this was possible to implement and emerged as a "best" match for these evaluations and organisations. How appropriate it is for IOs to embrace this approach for evaluation in general is an open question. This author would propose that it is, but evaluation fields, such as health and development evaluation may have a stronger influence and advocate for other approaches.

7.7. References

Alkin, M. C., & Taut, S. M. (2002). Unbundling evaluation use. *Studies in Educational Evaluation*, 29(1), 1-12.

Banathy, B. H. (2000). *Guided evolution of society: A systems view*. New York, NY: Kluwer Academic/Plenum Publishers.

Bertalanffy, L. (1969). *General Systems Theory: Foundations, Development, Application.* New York, NY: Braziller.

Braverman, M. T., & Arnold, M. E. (2008). An evaluator's balancing act: Making decisions about methodological rigor. In M. T. Braverman, M. Engle, M. E. Arnold & R. A. Rennekamp (Eds.), *Program evaluation in a complex organizational system: Lessons from Cooperative Extension*. *New Directions for Evaluation*, 120, 71–86.

Broom, G., & Dozier, D. (1990). *Using research in public relations*. Upper Saddle River, NJ: Prentice Hall.

Busby, J. W. (2007). Bono made Jesse Helms cry: Jubilee 2000, debt relief, and moral action in international politics. *International Studies Quarterly*, *51*(2), 247-275.

Ciarlo, J. A. (Ed.). (1981*). Utilizing evaluation: Concepts and measurement techniques.* Beverly Hills, CA: Sage Publications.

Coffman, J. (2002). *Public communication campaign evaluation: An environmental scan of challenges, criticisms, practice, and opportunities*. Cambridge, MA: Harvard Family Research Project.

Contandriopoulos, D., & Brousselle, A. (2012). Evaluation models and evaluation use. *Evaluation*, 18(1), 61-77.

Cousins, J. B., & Leithwood, K. A. (1986). Current empirical research in evaluation utilization. *Review of Educational Research*, (536), 331-364.

Diffenbach, J. (1982). Influence diagrams for complex strategic issues. *Strategic Management Journal*, *3*(2), 133-146.

Fitzpatrick, J. L. (2012). An introduction to context and its role in evaluation practice. *New Directions for Evaluation*, 2012(135), 7-24.

Fitzpatrick, J. L., Christie, C., & Mark, M. M. (2009). *Evaluation in action: Interviews with expert evaluators*. Thousand Oaks, CA: Sage Publications.

Grunig, J. (2008). Conceptualizing quantitative research in public relations. In B. van Ruler, A. Tkalac-Verčič & D. Verčič (Eds.), *Public relations metrics* (pp. 88-119). New York, NY: Routledge.

Henry, G. T., & Mark, M. M. (2003). Beyond Use: Understanding Evaluation's Influence on Attitudes and Actions. *American Journal of Evaluation*. Vol. 24, no. 3: 293-314.

Højlund, S. (2014). Evaluation use in evaluation systems–the case of the European Commission. *Evaluation*, 20(4), 428-446.

International Labour Office. (2013). *ILO policy guidelines for results-based evaluation: principles, rationale, planning and managing for evaluations*. International Labour Office, Evaluation Unit (EVAL), second edition, Geneva.

Jelen, A. (2008). The nature of scholarly endeavours in Public Relations. In B. van Ruler, A. Tkalac-Verčič & D. Verčič (Eds.), *Public relations metrics* (pp. 36-59). New York, NY: Routledge.

Johnson, K., Greenseid, L. O., Toal, S. A., King, J. A., Lawrenz, F., & Volkov, B. (2009). Research on Evaluation Use A Review of the Empirical Literature From 1986 to 2005. *American Journal of Evaluation*, 30(3), 377-410.

Lehmann, J. (2007). *Evaluating Mass Media Health Communication*. (Unpublished master's dissertation). University of Fribourg, Fribourg, Switzerland.

Mark, M.M., & Henry, G.T. (2004). The Mechanisms and Outcomes of Evaluation Influence. *Evaluation*. Vol. 10(1): 35–57.

Mayne, J. (2007). *Best Practices in Results-based Management: A Review of Experience: A Report for the United Nations Secretariat, Volume 1: Main Report., July 2007.* Retrieved from:
http://www.rocketsciencelab.co.uk/hipm/docs/john_mayne_un_1.pdf

Rogers, P. J. (2008). Using programme theory for complex and complicated programs. *Evaluation*, 14(1): 29-48.

Sanderson, I. (2000). Evaluation in complex policy systems. *Evaluation*, 6(4), 433-454.

Valente, T. (2001). Evaluating communication campaigns. In R. Rice & C. Atkins (Eds.), *Public communication campaigns* (pp. 105-124). Newbury Park, CA: Sage Publications.

Annexes

Annex 1: List of international organisations

The following list was established by this author in January 2011. The following abbreviations are used under "Type": INGO: international non-governmental organisation; IGO: intergovernmental organisation; IGO (UN): intergovernmental organisation (United Nations entity). A cross in the Included (Inc.) column indicates that this organisation was featured in this thesis (article 1, 2 or 3).

Name of International Organisation	*Type*	*Inc.*
AARP	INGO	
Abdus Salam International Centre for Theoretical Physics	IGO (UN)	X
Academic Council on the United Nations System	INGO	
ActionAid	INGO	X
Adventist Development and Relief Agency	INGO	
Africa-America Institute	INGO	
African Development Bank	IGO	
African Union	IGO	
African, Caribbean and Pacific Group of States	IGO	
Agence internationale pour le développement	INGO	
Asian Development Bank	IGO	
Asian Legal Resource Centre	INGO	
Asian-African Legal Consultative Organization	IGO	
Association for Progressive Communications	INGO	
Association for Women's Rights in Development	INGO	
Association of Caribbean States	IGO	
Association of Southeast Asian Nations	IGO	
AVSI Foundation	INGO	
Black Sea Economic Cooperation Organization	IGO	
CARE International	INGO	X
Caribbean Community	IGO	
Caritas Internationalis	INGO	
Central American Integration System	IGO	
Centre Europe-tiers monde	INGO	
Centro di Ricerca e Documentazione Febbraio 74	INGO	
CIDSE	INGO	
CIVICUS: World Alliance for Citizen Participation	INGO	
Commission of the Churches on International Affairs of the World Council of Churches	INGO	
Common Fund for Commodities	IGO	
Commonwealth of Independent States	IGO	
Commonwealth Secretariat	IGO	
Community of Portuguese-speaking Countries	IGO	
Community of Sahelo-Saharan States	IGO	
Comprehensive Nuclear-Test-Ban Treaty Organization	IGO (UN)	
Conference of NGOs in Consultative Relationship with the UN	INGO	

Consumers International	INGO	
Coordination SUD	INGO	
Council of Europe	IGO	
Customs Cooperation Council	IGO	
East African Community	IGO	
Economic and Social Commission for Asia and the Pacific	IGO (UN)	
Economic and Social Commission for Western Asia	IGO (UN)	
Economic Commission for Africa	IGO (UN)	
Economic Commission for Europe	IGO (UN)	
Economic Commission for Latin America and the Caribbean	IGO (UN)	
Economic Community of Central African States	IGO	
Economic Community of West African States	IGO	
Economic Cooperation Organization	IGO	
Energy Charter Conference	IGO	
Environmental Development Action in the Third World	INGO	
Eurasian Development Bank	IGO	
European Community	IGO	X
Food and Agriculture Organization of the United Nations	IGO (UN)	X
Foundation for the Social Promotion of Culture	INGO	
Friends World Committee for Consultation	INGO	
Good Neighbors International	INGO	
Green Cross International	INGO	
Greenpeace International	INGO	
Hague Conference on Private International Law	IGO	
HelpAge International	INGO	
Indian Ocean Commission	IGO	
InterAction: American Council for Voluntary International Action	INGO	
Inter-Agency Working Group on Evaluation	IGO (UN)	
Inter-American Development Bank	IGO	
International Alliance of Women	INGO	
International Association for Religious Freedom	INGO	
International Association of Lions Clubs	INGO	
International Association of Soldiers for Peace	INGO	
International Atomic Energy Agency	IGO (UN)	X
International Bank for Reconstruction and Development	IGO (UN)	
International Bureau of Education	IGO (UN)	
International Center for Research on Women	INGO	
International Centre for Genetic Engineering and Biotechnology	IGO (UN)	
International Centre for Migration Policy Development	IGO	
International Centre for Science and High Technology	IGO (UN)	
International Centre for Settlement of Investment Disputes	IGO (UN)	
International Chamber of Commerce	INGO	
International Civil Aviation Organization	IGO (UN)	
International Committee of the Red Cross	IGO	X
International Confederation of Free Trade Unions	INGO	
International Cooperative Alliance	INGO	
International Council for Adult Education	INGO	
International Council of Environmental Law	INGO	
International Council of Voluntary Agencies	INGO	

International Council of Women	INGO	
International Council on Social Welfare	INGO	
International Court of Justice	IGO (UN)	
International Criminal Court	IGO	
International Criminal Police Organization (INTERPOL)	IGO	
International Criminal Tribunal for Rwanda	IGO (UN)	
International Criminal Tribunal for the Former Yugoslavia	IGO (UN)	
International Development Law Organization	IGO	
International Federation of Agricultural Producers	INGO	
International Federation of Associations of the Elderly	INGO	
International Federation of Business and Professional Women	INGO	
International Federation of Red Cross and Red Crescent Societies	IGO	
International Federation of Settlements and Neighbourhood Centres	INGO	
International Federation on Ageing	INGO	
International Fund for Agricultural Development	IGO (UN)	
International Hydrographic Organization	IGO	
International Institute for Applied Systems Analysis	INGO	
International Institute for Democracy and Electoral Assistance	IGO	
International Institute for Non-Aligned Studies	INGO	
International Institute on Ageing	IGO (UN)	
International Labour Organization	IGO (UN)	X
International Maritime Organization	IGO (UN)	
International Monetary Fund	IGO (UN)	
International Movement ATD Fourth World	INGO	
International Organization for Migration	IGO	X
International Organization for Standardization	INGO	
International Organization of Employers	INGO	
International Organization of la Francophonie	IGO	
International Planned Parenthood Federation	INGO	X
International Research and Training Institute for the Advancement of Women	IGO (UN)	
International Save the Children Alliance	INGO	
International Seabed Authority	IGO	
International Seabed Authority	IGO (UN)	
International Social Security Association	INGO	
International Strategy for Disaster Reduction	IGO (UN)	
International Telecommunication Union	IGO (UN)	X
International Trade Centre 1	IGO (UN)	
International Trade Union Confederation	INGO	
International Tribunal for the Law of the Sea	IGO	
International Union for Conservation of Nature	IGO	X
Inter-Parliamentary Union	IGO	
Inter-Press Service International Association	INGO	
Islamic Development Bank Group	IGO	
Joint Inspection Unit	IGO (UN)	
Joint United Nations Programme on HIV/AIDS	IGO (UN)	
Latin American Economic System	IGO	
Latin American Integration Association	IGO	
Latin American Parliament	IGO	
League of Arab States	IGO	

Liberal International	INGO	
Médecins du monde (international)	INGO	
Médecins sans frontières (international)	INGO	
Muslim World League	INGO	
New Humanity	INGO	
Office for Outer Space Affairs	IGO (UN)	
Office for the Coordination of Humanitarian Affairs	IGO (UN)	X
OPEC Fund for International Development	IGO	
Orbicom: réseau des Chaires UNESCO en communication	INGO	
Organization for Economic Cooperation and Development	IGO	
Organization for Industrial, Spiritual and Cultural Advancement International	INGO	
Organization for Security and Cooperation in Europe	IGO	
Organization for the Prohibition of Chemical Weapons	IGO (UN)	
Organization of American States	IGO	
Organization of Eastern Caribbean States	IGO	
Organization of Islamic Capitals and Cities	INGO	
Organization of the Islamic Conference	IGO	
Organization of World Heritage Cities	INGO	
Oxfam International	INGO	X
Pacific Islands Forum	IGO	
Parliamentarians for Global Action	INGO	
Partners in Population and Development	IGO	
Permanent Court of Arbitration	IGO	
Regional Centre on Small Arms and Light Weapons in the Great Lakes Region	IGO	
ReliefWeb	IGO (UN)	
Rotary International	INGO	
Shanghai Cooperation Organization	IGO	
Socialist International	INGO	
Society for International Development	INGO	
Soroptimist International	INGO	
South Asian Association for Regional Cooperation	IGO	
South Centre	IGO	
Southern African Development Community	IGO	
Sovereign Military Order of Malta	IGO	
Transnational Radical Party	INGO	
United Cities and Local Governments	INGO	
United Nations Board of Auditors	IGO (UN)	
United Nations Capital Development Fund	IGO (UN)	
United Nations Children's Fund	IGO (UN)	X
United Nations Commission on International Trade Law	IGO (UN)	
United Nations Communications Group	IGO (UN)	
United Nations Conference on Trade and Development	IGO (UN)	X
United Nations Convention to Combat Desertification	IGO (UN)	
United Nations CyberSchoolBus	IGO (UN)	
United Nations Development Fund for Women	IGO (UN)	
United Nations Development Group	IGO (UN)	X
United Nations Development Programme	IGO (UN)	X
United Nations Educational, Scientific and Cultural Organization	IGO (UN)	X
United Nations Environment Programme	IGO (UN)	X

United Nations Framework Convention on Climate Change	IGO (UN)	
United Nations Fund for International Partnerships	IGO (UN)	
United Nations High Commissioner for Human Rights, Office of the	IGO (UN)	X
United Nations High Commissioner for Refugees, Office of the	IGO (UN)	X
United Nations Human Settlements Programme (HABITAT)	IGO (UN)	X
United Nations Industrial Development Organization	IGO (UN)	
United Nations Institute for Disarmament Research	IGO (UN)	
United Nations Institute for Training and Research	IGO (UN)	
United Nations Interregional Crime and Justice Research Institute	IGO (UN)	
United Nations Non-Governmental Liaison Service	IGO (UN)	
United Nations Office on Drugs and Crime	IGO (UN)	
United Nations Population Fund	IGO (UN)	
United Nations Postal Administration	IGO (UN)	
United Nations Relief and Works Agency for Palestine Refugees in the Near East	IGO (UN)	
United Nations Research Institute for Social Development	IGO (UN)	
United Nations Resident Coordinators Network	IGO (UN)	
United Nations Scientific Committee on the Effects of Atomic Radiation	IGO (UN)	
United Nations Volunteers	IGO (UN)	
United Nations World Tourism Organization	IGO (UN)	
Universal Postal Union	IGO (UN)	
Women's Federation for World Peace International	INGO	
Women's International Democratic Federation	INGO	
WomenWatch	IGO (UN)	
World Association of Girl Guides and Girl Scouts	INGO	
World Bank Group	IGO (UN)	X
World Blind Union	INGO	
World Confederation of Labour	INGO	
World Confederation of Productivity Science	INGO	
World Conference of Religions for Peace	INGO	
World Economic Forum	INGO	
World Family Organization	INGO	
World Federation of Democratic Youth	INGO	
World Federation of Trade Unions	INGO	
World Federation of United Nations Associations	INGO	
World Fellowship of Buddhists	INGO	
World Food Programme	IGO (UN)	
World Health Organization	IGO (UN)	
World Information Transfer	INGO	
World Intellectual Property Organization	IGO (UN)	
World Meteorological Organization	IGO (UN)	
World Movement of Mothers	INGO	
World Muslim Congress	INGO	
World Organization of the Scout Movement	INGO	
World Trade Organization	IGO (UN)	
World Veterans Federation	INGO	
World Vision International	INGO	
World Wide Fund for Nature International	INGO	
Zonta International	INGO	

Annex 2: Description: two cases of articles 2 and 3

This annex describes the two organisations and the communication campaigns that feature in articles 2 and 3 and summarises the research carried out with them.

The international Committee of the Red Cross (ICRC)

The ICRC is the founding body of the International Red Cross and Red Crescent (RCRC) Movement and has its headquarters in Geneva, Switzerland. The mandate of the ICRC focuses on protecting and assisting the victims of armed conflict. The ICRC operates in some 80 countries and employs 12,000 persons worldwide. The ICRC is considered to be an intergovernmental organisation.

The evaluation of the ICRC global campaign "Our world. Your move." on the 60th anniversary of the Geneva Conventions, which ran from January 2009 to December 2009, was a focus of article 2.

The broad aim of the campaign was to raise awareness of today's major humanitarian challenges and the work of the RCRC Movement. It also intended to encourage individuals to "make a move" for humanity based on the notion that simple gestures can make a difference. Three significant milestones for the RCRC Movement were marked in 2009:
• 150th anniversary of the Battle of Solferino;
• 90th anniversary of the founding of the International Federation of Red Cross and Red Crescent Societies;
•60th anniversary of the Geneva Conventions.

The key products and activities of the campaign included:
•Campaign logo, slogan and accompanying promotional material
•Events around the key dates of May (world Red Cross day), June (Battle of Solferino) and August (Geneva Conventions);

• Two photo exhibitions;

• Opinion research in eight countries;

• Joint media productions;

•Campaign portal and accompanying social media campaign;

• Street marketing in Geneva;

•Video clips, merchandise, promotional material and publications.

National Red Cross and Red Crescent Societies and field delegations of the ICRC and the International Federation undertook a wide variety of activities as part of the campaign.

This author carried out the evaluation of the campaign in collaboration with the communication unit of the ICRC, using multiple qualitative and quantitative methods. Article 2 provides an analysis and reflection of the challenges for communication evaluation based on this campaign evaluation and another evaluation carried out with the Office of the High Commissioner for Human Rights (OHCHR).

Some four years after the evaluation, this author returned to the ICRC to investigate the use of evaluation findings amongst communication staff. This study of use was mainly carried out through interviews with communication staff which were then analysed with a conceptual framework of use, as detailed in article 3.

The collaboration with the ICRC for this thesis ranged over some six years, from June 2008 to July 2014.

The Office of the High Commissioner for Human Rights (OHCHR)

The OHCHR is a component of the United Nations system and has its headquarters in Geneva, Switzerland. The mandate of OHCHR focuses on promoting and protecting

human rights. OHCHR operates in some 60 countries and employs 1,300 persons worldwide. The OHCHR is considered to be an intergovernmental organisation.

The evaluation of the OHCHR global campaign on the 60th anniversary of the Universal Declaration of Human Rights (UDHR), which ran from December 2007 to December 2008, was a focus of article 2.

The broad goal of the campaign "UDHR60" was to increase knowledge and awareness of human rights among the broadest audiences possible with an aim of empowering rights holders to claim and enjoy their rights.

UDHR60 focused on encouraging and guiding OHCHR partners − OHCHR country offices, the UN system, civil society, governments and national institutions, educational institutions, etc., to mark the anniversary in their own way. The strategy and production of some central resources and events was managed by OHCHR but UDHR60 was designed as a UN-wide campaign and not as a single agency initiative. Some of the key activities of UDHR60 managed by OHCHR included:

•Campaign logo and key materials (booklets, information kit, poster)

•An artistic project (a film, poster series and a book)

• Special events

• A media campaign

• A web campaign (specific website and section on OHCHR central website)

And partner-led UDHR60 activities included:

• Adaptation of campaign material

• Special events and conferences

• Grassroots mobilization

• Media campaign

• Web campaigns

 This author carried out the evaluation of the campaign in collaboration with the communication unit of the OHCHR, using multiple qualitative and quantitative methods.

Article 2 provides an analysis and reflection of the challenges for communication evaluation based on this campaign evaluation and the above-mentioned ICRC campaign.

Some four years after the evaluation, this author returned to the OHCHR to investigate the use of evaluation findings amongst the communication staff. As for the ICRC study of use, this was mainly carried out through interviews with communication staff which were then analysed with a conceptual framework of use, as detailed in article 3.

The collaboration with the OHCHR for this thesis ranged over some six years, from May 2008 to July 2014.

The figure on the next page maps out the interaction between the author and the organisations from 2008 to 2014. Of note, there is break in the horizontal X axis which indicates a gap of some four years where there was no contact between the author and the organisations (from 2010 to 2014).

Figure I: Interactions and milestones of the collaboration with the two organisations (2008-2014)

Notes for Figure I

One greyed square is equal to one interaction (physical meeting) with the organisation; the unit of time is months with the vertical lines separating each month.

Annex 3: Consolidated Bibliography

Albaek, E. (1995). Between knowledge and power: Utilization of social science in public policy making. *Policy Sciences*, 28(1), 79-100.

Alkin, M. C. (Ed.). (2012). *Evaluation Roots: A Wider Perspective of Theorists' Views and Influences*. London: Sage Publications.

Alkin, M. C., Daillak, R., & White, P. (1979). *Using evaluations: Does evaluation make a difference.* Beverly Hills, CA: Sage Publications.

Alkin, M. C., & Taut, S. M. (2002). Unbundling evaluation use. *Studies in Educational Evaluation*, 29(1), 1-12.

Alvarez, J. (2006). International Organisations: Then and Now. *The American Journal of International Law, 100*, (2), 324-347.

Amagoh, F. (2008). Perspectives on organizational change: systems and complexity theories. *The Innovation Journal: The public sector innovation journal*, 13(3), 1-14.

AMEC. (2010). *Barcelona Declaration of Measurement Principles.* Paper presented at: the 2nd European Summit on Measurement, Barcelona, 17-20 June 2010 [PDF document]. Retrieved from:
http://www.instituteforpr.org/wp-content/uploads/BarcelonaPrinciplesOct2010.pdf

Ayers, T. D. (1987). Stakeholders as partners in evaluation: A stakeholder-collaborative approach. *Evaluation and Program Planning*, 10, 263-271.

Bamberger, M., Rugh, J., & Mabry, L. (2006). *Real world evaluation: Working under budget, time, data and political constraints*. Thousand Oaks, CA: Sage Publications.

Banathy, B. H. (2000). *Guided evolution of society: A systems view*. New York, NY: Kluwer Academic/Plenum Publishers.

Banks, S. P. (1995). *Multicultural public relations: A social-interpretive approach*. Thousand Oaks, CA: Sage Publications.

Barnett, M., & Finnemore, M. (1999). The Politics, Power, and Pathologies of International Organisations. *International Organisation*, *53*(4), 699-732.

Bauer, M. W. (2000). Classical content analysis: a review. In M. W. Bauer & G. Gaskell (Eds.), *Qualitative analysis with text, image and sound* (pp. 131-251). London: Sage Publications.

Beach, D., & Pedersen, R. B. (2013). *Process-Tracing Methods: Foundations and Guidelines*. Ann Arbor: University of Michigan Press.

Bertalanffy, L. (1969). *General Systems Theory: Foundations, Development, Application.* New York, NY: Braziller.

Bertrand, J. T., O'Reilly, K., Denison, J., Anhang, R., & Sweat, M. (2006). Systematic review of the effectiveness of mass communication programs to change HIV/AIDS-related behaviors in developing countries. *Health Education Research*, *21*, 567–597.

Bhatt, G. D. (2001). Knowledge management in organizations: examining the interaction between technologies, techniques, and people. *Journal of knowledge management*, 5(1), 68-75.

Birkland, T. A. (2014). *An introduction to the policy process: Theories, concepts and models of public policy making*. New York, NY: Routledge.

Braverman, M. T., & Arnold, M. E. (2008). An evaluator's balancing act: Making decisions about methodological rigor. In M. T. Braverman, M. Engle, M. E. Arnold & R. A. Rennekamp (Eds.), *Program evaluation in a complex organizational system: Lessons from Cooperative Extension*. *New Directions for Evaluation*, 120, 71–86.

Broom, G. M., & Dozier, D. M. (1990). *Using research in public relations: Applications to program management*. Englewood Cliffs, NJ: Prentice-Hall.

Brown, L. D., Ebrahim, A., & Batliwala, S. (2012). Governing international advocacy NGOs. *World Development*, 40(6), 1098-1108.

Cabanera-Verzosa, C. (2003). *Strategic Communication for Development Projects*. World Bank. Retrieved from:
http://siteresources.worldbank.org/EXTDEVCOMMENG/Resources/toolkitwebjan200 4.pdf

Campbell, D. T. (1984). Can we be scientific in applied social science. *Evaluation studies review annual*, Volume 9. Beverly Hills, CA: Sage Publications.

Caracelli, V. J. (2000). Evaluation use at the threshold of the twenty‑first century. *New directions for evaluation*, 2000(88), 99-111.

Carden, F., Smutylo, T., & Earl, S. (2001). *Outcome mapping: Building learning and reflection into development programs*. International Development Research Centre.

Charnovitz, S. (2006). Nongovernmental Organisations and International Law. *The American Journal of International Law*, *100*(2), 348-372.

Chartered Institute for Public Relations. (2004). *Best Practice in the Measurement and Reporting of Public Relations and ROI*. London: Chartered Institute for Public Relations.

Chen, H. (2005). *Practical Program Evaluation: Assessing and Improving Planning, Implementation and Effectiveness*. Thousand Oaks, CA: Sage Publications.

Chesterton, P. (2004). *Evaluation of the Meena Communication Initiative*, UNICEF Regional Office for South Asia, Kathmandu.

Christie, C. A. (2007). Reported influence of evaluation data on decision makers' actions: An empirical examination. *American Journal of Evaluation*, 28, 8-25.

Christie, C., & Fierro, L. (2012). Evaluation policy to implementation: An evaluation of scientifically based research in practice. *Studies in Educational Evaluation*, 38: 65–72.

Ciarlo, J. A. (Ed.). (1981*). Utilizing evaluation: Concepts and measurement techniques.* Beverly Hills, CA: Sage Publications.

Clarke, A. (1999). *Evaluation research: An introduction to principles, methods and practice*. London: Sage Publications.

Clay, E. J., & Schaffer, B. B (Eds.) (1986). *Room for Manoeuvre, An Explanation of Public Policy in Agriculture and Rural Development*. London: Heinemann.

Coffman, J. (2002). *Pubic communication campaign evaluation - An environmental scan of challenges, criticisms, practice and opportunities*. Harvard Family Research Project. Retrieved from: http://www.mediaevaluationproject.org/HFRP.pdf

Coffman, J. (2003). *Lessons in evaluating communications campaigns: Five case studies*. Harvard Family Research Project. Retrieved from: http://www.hfrp.org/content/download/1114/48619/file/lessons_comm_campaigns.pdf

Cohen, J. A. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37–46.

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: a new perspective on learning and innovation. *Administrative science quarterly*, 128-152.

Communications Consortium Media Center. (2004). *Guidelines for Evaluating Nonprofit Communications Efforts*. Retrieved from:
http://www.researchtoaction.org/2013/03/guidelines-for-evaluating-nonprofit-communications-efforts/

Connolly, P., & York, P. (2002). Evaluating capacity-building efforts for nonprofit organizations. *OD Practitioner*, *34*(4), 33-39.

Contandriopoulos, D., & Brousselle, A. (2012). Evaluation models and evaluation use. *Evaluation*, 18(1), 61-77.

Cook, T. D., & Campbell D. T. (1979). *Quasi Experimentation: Design and Analytical Issues for Field Settings.* Chicago, IL: Rand McNally.

Cousins, J. B. (1995). Assessing program needs using participatory evaluation: A comparison of high and marginal success cases. In J.B. Cousins & L. M. Earl (Eds.), *Participatory evaluation in education: Studies in evaluation use and organizational learning* (pp. 55-71). London: Falmer.

Cousins, J., & Chouinard, J. (2012). *Participatory evaluation up close: A review and integration of research-based knowledge*. Charlotte, NC: Information Age Press.

Cousins, J. B., Goh, S., Clark, S., & Lee, L. (2004). Integrating evaluative inquiry into the organizational culture: A review and synthesis of the knowledge base. *Canadian Journal of Program Evaluation*, 19(2), 99-141.

Cousins, J. B., Goh, S. C., Elliott, C. J., & Bourgeois, I. (2014). Framing the capacity to do and use evaluation. In J.B. Cousins, & I. Bourgeois (Eds.), *Organizational capacity to do and use evaluation. New Directions for Evaluation*, 141, 7–23.

Cousins, J., & Leithwood, K. (1986). Current Empirical Research on Evaluation Utilization. *Review of Educational Research*, Vol. 56, No. 3: 331-364.

Culbertson, H. M. (1996). Introduction. In H. M. Culbertson & N. Chen (Eds.), *International public relations: A comparative analysis* (pp. 1–13). Mawah, NJ: Lawrence Erlbaum.

Cutler, A. (2004). Methodical failure: The use of case study method by public relations researchers. *Public Relations Review*, 30(3), 365-375.

Cutlip, S. M., Center, A. H., & Broom, G. M. (2006). *Effective public relations* (9th ed.). Upper Saddle River, NJ: Prentice Hall.

Dahler-Larsen, P. (2012). *The Evaluation Society*. Paolo Alta, CA: Stanford Press.

Datta, L. (1994). Paradigm wars: A basis for peaceful coexistence and beyond. *New Directions for Program Evaluatio*n, 61: 53-70.

Davidson, J. (2005). *Evaluation methodology basics: the nuts and bolts of sound evaluation.* London: Sage Publications.

de Certeau, M. (1990). *L'invention du quotidien. 1. arts de faire* (1ère édition 1980). Paris: Gallimard.

Deibert, R. (2000). International Plug n' Play? Citizen Activism, the Internet, and Global Public Policy. *International Studies Perspectives*, Vol. 1, No. 3: 255-272.

DeMars, W. E. (2005). *NGOs and Transnational Networks: Wild Cards in World Politics*. London: Pluto.

Dettmer, H. W. (2006). Systems and constraints: the concept of leverage. Port Angeles, WA: *Goal Systems International.* Retrieved January, 21, 2009.Retrieved from: http://www.goalsys.com/systemsthinking/documents/Part-6-SystemsandConstraints.pdf

Diffenbach, J. (1982). Influence diagrams for complex strategic issues. *Strategic Management Journal*, *3*(2), 133-146.

Dijkzeul, D., & Beigbeder, Y. (Eds.). (2003). *Rethinking international organizations: pathology and promise*. New York, NY: Berghahn Books.

Donaldson, S.I., & Lipsey, M. W. (2006). Roles for theory in contemporary evaluation practice: Developing practical knowledge. In I. Shaw, J. Greene & M. Mark, *SAGE Handbook of Evaluation: Policies, Programs and Practices (pp. 56-75),* London: Sage Publications.

Dozier, D. (1990). The innovation of research in public relations practice: Review of a program of studies. *Public Relations Research Annual*, 2, 3-28.

Dozier, D., & Broom, G. (1995). Evolution of the manager role in public relations practice. *Journal of Public Relations Research, 7*(1), 3-26.

Dozier, D., & Ehling, W. (1992) Evaluation of PR Programs: What the Literature Tells Us About Their Effects. In J. Grunig (Ed.), *Excellence in Public Relations & Communication Management* (pp. 159-184). Hillsdale, NJ: Lawrence Erlbaum Associates.

Dozier, D., Grunig, L., & Grunig, J. (2001). Public Relations as Communications Campaign. In R. Rice & C. Atkins (Eds.), *Public communication campaigns* (3rd ed.) (pp. 231-248). Newbury Park, CA: Sage Publications.

Drazin, R., & Van de Ven, A. H. (1985). Alternative forms of fit in contingency theory. *Administrative science quarterly*, 514-539.

Dubey, D. & Bardhan, A. (1981). *Evaluation of communication programmes*. UNESCO Communication Research Module No. 5. New Delhi: Caxton Press.

Dubin, R. (1976). Theory building in applied areas. In M.D. Dunnette (Ed.), *Handbook of Industrial/organizational psychology* (pp. 17-39). New York, NY: Rand McNally.

Duncan, B. (2002). Health policy in the European Union: how it's made and how to influence it. *British Medical Journal*, 324(7344): 1027–1030.

Dwyer, S. C., & Buckle, J.L. (2009). The Space Between: On Being an Insider-Outsider in Qualitative Research. *International Journal of Qualitative Methods* 8:54–63.

Elliott, J., Holland, J., & Thomson, R. (2008). Longitudinal and Panel Studies. In P. Alasuutari, L. Bickman & J. Brannen, *The SAGE Handbook of Social Research Methods* (pp. 228–249). Thousand Oaks, CA: Sage Publications.

Elman, C., & Elman, M. (Eds.). (2003). *Progress in International Relations Theory*. Cambridge, MA: MIT Press.

European Commission. (2000). *Focus on Results: Strengthening evaluation of Commission activities, SEC (2000)1051 - 26/07/2000, Communication to the Commission from Mrs. Schreyer in agreement with Mr. Kinnock and the President, according to action 16 of the Action Plan for Reform.* Retrieved from:

http://ec.europa.eu/dgs/information_society/evaluation/data/pdf/lib_master/sec2000_1051
_strengthening_eval.pdf.

Fairchild, M. (2003). *IPR toolkit: Media Evaluation Edition*. London: Charted Institute for Public Relations.

Fischer, R. (1995). Control construct design in evaluating campaigns. *Public Relations Review*, *21*(1), 45-58.

Fitz-Gibbon, C. T., & Morris, L. L. (1996). Theory-based evaluation. *Evaluation Practice*, 17(2), 177-184.

Fitzpatrick, J. L. (2012). An introduction to context and its role in evaluation practice. *New Directions for Evaluation*, 2012(135), 7-24.

Fitzpatrick, J. L., Christie, C., & Mark, M. M. (2009). *Evaluation in action: Interviews with expert evaluators.* Thousand Oaks, CA: Sage Publications.

Flay, B., & Cook, T. (1981). Evaluation of Mass Media Prevention Campaigns. In R. Rice & W. Paisley, *Public Communication Campaigns* (pp. 239–264). London: Sage Publications.

Fonbrum, C. (1996). *Reputation: Realizing Value from the Corporate Image.* Boston: Harvard Business School Press.

Funnell, S., & Rogers, P. (2011) *Purposeful Programme Theory. Effective use of theories of change and logic models*. San Francisco: Jossey-Bass.

Gallagher, M. (1985). *Becoming aware: human rights and the family – a study based on four communication campaigns*. Paris: UNESCO.

Gaunt, R., & Wright, D. (2004). *PR Measurement*. London: Benchpoint.

Gillespie, A., & Cornish, F. (2014). Sensitizing questions: A method to facilitate analyzing the meaning of an utterance. *Integrative Psychological and Behavioral Science*, 48(4), 435-452.

Gonçalves B., & Ramasco J. (2008). Human dynamics revealed through Web analytics. *Physical Review*, 78, 1-7.

Goodman, R., & Jinks, D. (2003). Measuring the Effects of Human Rights Treaties. *European Journal of International Law*, 14(1), 171-183.

Greene, J. C., & Benjamin, L. (2001). The Merits of Mixing Methods in Evaluation. *Evaluation, 7*(1), 25-44.

Greene, J. C., Mark, M. M., & Shaw, I. (2006). Handbook of evaluation: policies, programs and practices. London: Sage publications.

Greene, J. G. (1988). Stakeholder participation and utilization in program evaluation. *Evaluation Review*, (12), 912-116.

Gregory, A. (2001). Public relations and evaluation: Does the reality match the rhetoric? *Journal of Marketing Communication*, 7:3: 171–89.

Gregory, A., Morgan, L., & Kelly, D. (2005). *Patterns of PR in Britain's 'most admired' companies and public sector organisations*. Leeds: Leeds Business School for Eloqui Public Relations.

Gregory, A., & Watson, T. (2008). Defining the gap between research and practice in public relations programme evaluation - towards a new research agenda. *Journal of Marketing Communications*, 14:5: 337-350.

Gregory, A., & White, J. (2008). Introducing the Chartered Institute of Public Relations' initiative. In B. van Ruler, A. Tkalac-Verčič & D. Verčič (Eds.), *Public relations metrics* (pp. 307–17). New York, NY: Routledge.

Grunig, J. (ed.) (1992). *Excellence in Public Relations and Communications Management.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Grunig, J. (2006). Furnishing the edifice: Ongoing research on public relations as a strategic management function. *Journal of Public Relations Research*, 18(2), 151–176.

Grunig, J. (2008). Conceptualizing quantitative research in public relations. In B. van Ruler, A. Tkalac-Verčič & D. Verčič (Eds.), *Public relations metrics* (pp. 88-119). New York, NY: Routledge.

Grunig, J. (2009). Paradigms of global public relations in an age of digitalisation. *PRism,* 6(2). Retrieved from:
http://www.prismjournal.org/fileadmin/Praxis/Files/globalPR/GRUNIG.pdf

Grunig, J., & Hunt, T. (1984). *Managing Public Relations*. Fort Worth, TX: HBJ Publishers.

Guba, E., & Lincoln, Y. (1989). *Fourth generation evaluation.* London: Sage Publications.

Habermas, J. (1974). *Theory and practice*. Boston, MA: Beacon Press.

Habicht, J. P., Victora, C. G., & Vaughan, J. P. (1999). Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. *International journal of epidemiology*, *28*(1), 10-18.

Hall, M. (2008). The effect of comprehensive performance measurement systems on role clarity, psychological empowerment and managerial performance. *Accounting, Organizations and Society,* 33 (2-3), 141-163.

Hall, M. (2011). Do comprehensive performance measurement systems help or hinder managers' mental model development? *Management Accounting Research*, 22 (2), 68-83.

Harden, A., & Thomas J. (2005). Methodological Issues in Combining Diverse Study Types in Systematic Reviews. *International Journal of Social Research Methodology, 8*, 257–271.

Henry, G. T., & Gordon, C. S. (2003). Driving less for better air: Impacts of a public information campaign. *Journal of Policy Analysis and Management*, 22: 45–63.

Henry, G. T., & Mark, M. M. (2003). Beyond Use: Understanding Evaluation's Influence on Attitudes and Actions. *American Journal of Evaluation*. Vol. 24, no. 3: 293-314.

Højlund, S. (2014a). Evaluation use in the organizational context–changing focus to improve theory. *Evaluation*, 20(1), 26-43.

Højlund, S. (2014b). Evaluation use in evaluation systems–the case of the European Commission. *Evaluation*, 20(4), 428-446.

Holtzhausen, D., & Zerfass, A. (Eds.). (2014). *The Routledge handbook of strategic communication*. Chicago, IL: Routledge.

Hon, L. (1997). What Have You Done For Me Lately? Exploring Effectiveness in Public Relations. *Journal of Public Relations Research*, 9(1), 1-30.

Hon, L. (1998). Demonstrating Effectiveness in Public Relations: Goals, Objectives, and Evaluation. *Journal of Public Relations Research*, 10, 103–135.

Hon, L., & Grunig, J. (1999). *Guidelines for measuring relationships in public relations*. Gainesville, FL: Institute for Public Relations. Retrieved from: http://www.instituteforpr.org/topics/measuring-relationships/

Hornik, R. (2002). *Public Health Communication. Evidence for behaviour change.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Huhn, J. Sass, J., & Storck, C. (2011). *Position paper - Communication Controlling How to maximize and demonstrate the value creation through communication*. Berlin: German Public Relations Association (DPRG). Retrieved from: http://ec.europa.eu/dgs/information_society/evaluation/data/pdf/lib_master/sec2000_1051 _strengthening_eval.pdf

Hutchings, C., & Bowman, K. (2013). Lessons learned from a year of advocacy impact evaluation. *Oxfam Policy and Practice Blog*. Retrieved from: http://policy-practice.oxfam.org.uk/blog/2013/02/lessons-learned-from-a-year-of-advocacy-impact-evaluation

Hwang, Y. (2012). Social Diffusion of Campaign Effects: Campaign-Generated Interpersonal Communication as a Mediator of Anti-tobacco Campaign Effects. *Communication Research*, 39 (1), 120-141.

International Committee of the Red Cross. (2010). *About the International Committee of the Red Cross.* Retrieved from: http://www.icrc.org/eng/who-we-are/overview-who-we-are.htm

International Labour Office. (2013). *ILO policy guidelines for results-based evaluation: principles, rationale, planning and managing for evaluations*. International Labour Office, Evaluation Unit (EVAL), second edition, Geneva.

International Organisation for Migration. (2002). *Awareness Raising and Information Strategy on People in Need of International Protection - Evaluation Report*.

Jelen, A. (2008). The nature of scholarly endeavours in Public Relations. In B. van Ruler, A. Tkalac-Verčič & D. Verčič (Eds.), *Public relations metrics* (pp. 36-59). New York, NY: Routledge.

Johnson, K., Greenseid, L. O., Toal, S. A., King, J. A., Lawrenz, F., & Volkov, B. (2009). Research on Evaluation Use - A Review of the Empirical Literature From 1986 to 2005. *American Journal of Evaluation*, 30(3), 377-410.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness* (No. 6). Chicago, IL: Harvard University Press.

Julnes, P. D. L., & Holzer, M. (2001). Promoting the utilization of performance measures in public organizations: An empirical study of factors affecting adoption and implementation. *Public Administration Review*, 61(6), 693-708.

Juma, C., & Clark, N. (1995). Policy research in sub-Saharan Africa: An Exploration. *Public Administration and Development*, Vol. 15, 121–137.

Kaldor, M., Moore, H., & Selchow, S. (Eds.). (2012). *Global Civil Society 2012: Ten Years of Critical Reflection.* Basingstoke: Palgrave Macmillan.

Karpf, D. (2010). Online Political Mobilization from the Advocacy Group's Perspective: Looking Beyond Clicktivism. *Policy & Internet,* 2 (4), 7-41.

Katz, D., & Kahn, R. L. (1966). *The social psychology of organizations*. New York, NY: Wiley.

Keck, M., & Sikkink, K. (1998). *Activists Beyond Borders: Advocacy Networks in international politics*. Ithaca, NY: Cornell.

Kelly, K. S. (2001). Importance and use of formal research and evaluation. In R. L. Heath (Ed.), *Handbook of public relations* (pp. 279-290). Thousand Oaks, CA: Sage Publications.

Kennedy, M., & Abbatangelo, M. (2005). *Guidance for Evaluating Mass Communication Health Initiatives: Summary of an Expert Panel Discussion* (held May 3-4, 2004). Centers for Disease Control and Prevention, Office of Communication, Atlanta, USA.

Keohanea, R., & Nye, J. (1974). Transgovernmental Relations and International Organizations. *World Politics*, 27:01: 39-62.

Kerstetter, K. (2012). Insider, outsider, or somewhere in between: The impact of researchers' identities on the community-based research process. *Journal of Rural Social Sciences*, 27(2), 99–117.

Kickbusch, I. (2003).The Contribution of the World Health Organization to a New Public Health and Health Promotion. *American Journal of Public Health*, Vol. 93, No. 3: 383-388.

Kirkhart, K. (2000). Reconceptualizing evaluation use: An integrated theory of influence. *New directions for evaluation 2000*, no. 88 (2000): 5-23.

Lawson, R. Hatch, T., & Desroches, D. (2007). *Scorecard best practices: design, implementation, and evaluation*. Hoboken, NJ: John Wiley and Sons.

Lazarsfeld, P. (1948). Communication research and the social psychologist. In W. Dennis, *Current Trends in Social Psychology* (pp. 218-273). Pittsburg, PA: University of Pittsburgh Press.

Leeuw, F.L., & Furubo, J-E. (2008). Evaluation systems: what are they and why study them? *Evaluation,* 14: 157–69.

Lehmann, I. (1999). *Peacekeeping and Public Information, Caught in the crossfire.* London: Frank Cass.

Lehmann, J. (2007). *Evaluating Mass Media Health Communication.* (Unpublished master's dissertation). University of Fribourg, Fribourg, Switzerland.

Leinemann, R., & Baikaltseva, E. (2004). *Media relations measurement: determining the value of PR to your company's success.* London: Gower Publishing.

Lennie, J., & Tacchi, J. (2011). *Researching, monitoring and evaluating communication for development: Trends, challenges and approached.* Report on a literature review and Consultations with Expert Reference Group and UN Focal Points on C4D. United Nations Inter-Agency Group on Communications for Development, New York, NY: UNICEF. Retrieved from:
http://www.unicef.org/cbsc/files/RME-RP-Evaluating_C4D_Trends_Challenges__Approaches_Final-2011.pdf

Lennie, J., & Tacchi, J. (2013). *Evaluating Communication for Development: A Framework for Social Change.* New York, NY: Routledge.

Leviton, L. C., & Hughes, E. F. (1981). Research On the Utilization of Evaluations A Review and Synthesis. *Evaluation review*, 5(4), 525-548.

Levy, C. (1992). Gender and the environment: the challenge of cross-cutting issues in development policy and planning. *Environment and Urbanization*. 4.1: 134-149.

Lewin, K. (1948). 'Group decision and social change', reprinted in M. Gold (Ed.) (1999) *The Complete Social Scientist: A Kurt Lewin Reader.* Washington, DC: American Psychological Association. pp. 265–84.

Lieberman, E. (2005). Nested Analysis as a Mixed-Method Strategy for Comparative Research. *American Political Science Review,* 99(2): 435-452.

Likely, F., & Watson, T. (2013). Measuring the edifice: public relations measurement and evaluation practices over the course of 40 years. In K. Sriramesh, A. Zerfass & J-N. Kim (Eds.), *Public Relations and Communication Management: Current Trends and Emerging Topics* (pp. 143-162). New York, NY: Routledge.

Lincoln, Y. S., & Guba, E. (1985). *Naturalistic inquiry.* London: Sage Publications.

Lindenmann, W. (1993). An 'effectiveness yardstick' to measure public relations success. *Public Relations Quarterly, 38*, 7–10.

 Lindenmann, W. (2003). *Guidelines and Standards for Measuring the Effectiveness of PR Programs and Activities*. Institute for Public Relations. Retrieved from:
http://www.instituteforpr.org/topics/effectiveness-programs-activities/

Lindenmann, W. (2006). *Public Relations Research for Planning and Evaluation.* Gainesville, FL: Institute for Public Relations. Retrieved from:
http://www.instituteforpr.org/wp-content/uploads/2006_Planning_Eval2.pdf

Luhmann, N. (1990). *Essays on self-reference*. New York, NY: Columbia University Press.

Luo, H. (2010). The Role for an Evaluator: A Fundamental Issue for Evaluation of Education and Social Programs. *International Education Studies*, *3*(2), P42.

Macnamara, J. (1992). Evaluation: The Achilles' heel of the PR profession. *International Public Relations (IPRA) Review,* 15, No. 2: 17–31.

Macnamara, J. (2005). *Jim Macnamara's Public Relations Handbook*. Sydney: Archipelago Press.

Macnamara, J. (2005). Media content analysis: Its uses, benefits and best practice methodology. *Asia Pacific Public Relations Journal*, 6(1), 1–34.

Macnamara, J. (2006). Two-tier evaluation can help corporate communicators gain management support. *Prism* (online PR Journal), Vol. 4, Issue 2. Retrieved from: http://www.prismjournal.org/fileadmin/Praxis/Files/Journal_Files/Evaluation_Issue/COMMENTARY_MACNAMARA.pdf

Macnamara, J. (2014). The 'toe bone to the head bone' logic model to connect public relations and corporate communication to organisation outcomes. *Prism* (online PR Journal), Vol. 11, Issue 1. Retrieved from: http://www.prismjournal.org/fileadmin/11_1/Macnamara_2014.pdf

Manheim, J. (2011). *Strategy in information and influence campaigns: how policy advocates, social movements, insurgent groups, corporations, governments, and others get what they want*. New York, NY: Routledge.

Mannell, J. (2014). Adopting, manipulating, transforming: Tactics used by gender practitioners in South African NGOs to translate international gender policies into local practice. *Health & place*, *30*, 4-12.

Mark, M. M., Cooksy, L. J., & Trochim, W. M. (2009). Evaluation policy: An introduction and overview. *New Directions for Evaluation*, 2009(123), 3-11.

Mark, M., & Henry, G. (2003). Beyond Use: Understanding Evaluation's Influence on Attitudes and Actions. *American Journal of Evaluation*. Vol. 24, no. 3: 293-314.

Mark, M., & Henry, G. (2004). The Mechanisms and Outcomes of Evaluation Influence. *Evaluation*. Vol. 10(1): 35–57.

Martens, K. (2002). Mission Impossible? Defining Nongovernmental Organisations. Voluntas*: International Journal of Voluntary and Nonprofit Organisations*, Vol. 13, No. 3: 271-285.

Mason, J. (2002). *Researching your own practice: The discipline of noticing*. London: Routledge.

Mathiason, J. (2011). *The craft of results-based management*. Retrieved from: http://sites.maxwell.syr.edu/intleval/Sessions/1-the_craft_of_evaluation.html

Mathison, S. (1994). Rethinking the evaluator role: Partnerships between organizations and evaluators. *Evaluation and Program Planning*, 17(3), 299-304.

Mathison, S. (2004). *Encyclopedia of evaluation.* London: Sage Publications.

Maxwell, J. A. (2012). *Qualitative research design: An interactive approach* (Vol. 41). Newbury Park, CA: Sage Publications.

Mayne, J. (2007). *Best Practices in Results-based Management: A Review of Experience: A Report for the United Nations Secretariat, Volume 1: Main Report., July 2007.* Retrieved from:http://www.rocketsciencelab.co.uk/hipm/docs/john_mayne_un_1.pdf

McIver, P., & Carmines, E. G. (1981). *Unidimensional scaling* (Sage University Paper Series on Quantitative Applications in the Social Science, series no. 24). Newbury Park, CA: Sage Publications.

McLean, A. (2000). From international organisations to international organisations. In T. Salmon & M. Imber (Eds.), *Issues in International Relations* (pp. 159-188). London: Routledge.

McNiff, J., & Whitehead, J. (2011). *All you need to know about action research.* London: Sage Publications.

Medina-Borja, A., & Triantis, K. (2007). A conceptual framework to evaluate performance of non-profit social service organisations. *International Journal of Technology Management*, 37 (1-2), 147-161.

Mendelsohn, H. (1973). Some reasons why information campaigns can succeed. *The Public Opinion Quarterly*, *37*(1), 50-61.

Meredith, J. (1992). Theory Building through Conceptual Methods. *International Journal of Operations and Production Management,* 13 (5): 3-11.

Michaelson, D., & Macleod, S. (2007). The Application of "Best Practices" In Public Relations Measurement and Evaluation Systems, *Public Relations Journal,* 1(1) 1-14.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded source book* (2nd ed.). Newbury Park, CA: Sage Publications.

Mingst, K. (2004). *Essentials of International Relations*, 3rd ed. New York, NY: W. W. Norton and Company.

Mintzberg, H., & Waters, J. A. (1985). Of strategies, deliberate and emergent. *Strategic Management Journal*, 6, 257–272.

Mohammed, S., & Ringseis, E. (2001). Cognitive diversity and consensus in group decision making: The role of inputs, processes, and outcomes. *Organizational Behavior and Human Decision Processes,* 85(2), 310-335.

Morozov, E. (2009). The Brave New World of Slacktivism, *Foreign Policy*, May 19, 2009. Retrieved from:
http://neteffect.foreignpolicy.com/posts/2009/09/05/from_slacktivism_to_activism?wpisrc=obinsite

Movius, L. (2010). Cultural Globalisation and Challenges to Traditional Communication Theories, *PLATFORM: Journal of Media and Communication*, 2(1): 6-18.

Mowbray, C. T. (1992). The role of evaluation in restructuring of the public mental health system. *Evaluation and Program Planning*, 15, 403-415.

Noar, S., Palmgreen, P., Chabot, M., Dobransky, N., & Zimmerman, R. (2009). 10-Year Systematic Review of HIV/AIDS Mass Communication Campaigns: Have We Made Progress? *Journal of Health Communication*, Vol. 14:15–42.

Office of the High Commissioner for Human Rights. (2009). *OHCHR   Report: Activities and Results.* Retrieved from:
http://www.ohchr.org/Documents/Press/OHCHR_Report_2008.pdf

Office of the High Commissioner for Human Rights. (2012). *OHCHR Report 2011*. Retrieved from:
http://www2.ohchr.org/english/ohchrreport2011/web_version/ohchr_report2011_web/allegati/3_About_OHCHR.pdf

O'Neil, G. (2013). Evaluation of international and non-governmental organizations' communication activities: A 15 year systematic review. *Public Relations Review,* 39, 572-574.

O'Neil, G. (2015). Implementing communication evaluation methodologies for two international organisations. *Submitted for publication.*

O'Neil, G., & Bauer, M.W. (2015). The use of communication evaluation findings in two international organisations. *Submitted for publication.*

Otero, E., & Cugelman, B. (2010). *Evaluation of Oxfam GB's Climate Change Campaign.* Oxfam GB.

Patton, M. Q. (2000). Utilization-Focused Evaluation. In D. Stufflebeam, G. Madaus & T. Kellaghan (Eds.), *Evaluation Models: Viewpoints on Educational and Human Services Evaluation* (pp. 425-538). Norwell, MA: Springer.

Patton, M. Q. (2008). Advocacy Impact Evaluation. *Journal of MultiDisciplinary Evaluation*, Vol. 5, no. 9.

Patton, M. Q. (2008). *Utilization-focused Evaluation.* London: Sage Publications.

Patton, M. Q. (2011). *Developmental evaluation: Applying complexity concepts to enhance innovation and use.* New York, NY: Guilford Press.

Patton, M. Q., Grimes, P. S., Guthrie, K. M., Brennan, N. J., French, B. D., & Blyth, D. A. (1977). In search of impact: An analysis of the utilization of federal health evaluation research. In C. H. Weiss (Ed.), *Using social research in public policy making* (pp. 141-163). New York, NY: D. C. Heath and Company.

Pawson, R., & Tilley, N. (1997). *Realistic Evaluation*, London: Sage Publications.

Perrin, B. (2006). *Moving from Outputs to Outcomes: Practical advice from governments around the world*. Managing for performance and results series, Washington, DC: World Bank. Retrieved from:

http://siteresources.worldbank.org/CDFINTRANET/Resources/PerrinReport.pdf

Petticrew, M., & Roberts, H. (2006). *Systematic Reviews in the Social Sciences: a Practical Guide*. Oxford: Blackwell Publishing.

Picciotto, R. (2012). Experimentalism and development evaluation: Will the bubble burst? *Evaluation*, 18 (2): 213-229.

Pohl, G., & Vanderventer, D. (2001). The workplace, undergraduate education and career preparation. In R.L. Heath (Ed.), *Handbook of public relations* (pp. 357–68). Newbury Park, CA: Sage Publications.

Porto, M. (2007). Fighting AIDS Among Adolescent Women: Effects of a Public Communication Campaign in Brazil, *Journal of Health Communication,* 12, 2, 121-133

Radnor, H. (2001). *Researching your professional practice: Doing interpretive research*. Buckingham: Open University Press.

Rantanen, T. (2005). *The Media and Globalization*. London: Sage Publications.

Reason, P., & Bradbury, H. (2001) *Handbook of Action Research*. London: Sage Publications.

Reeves, T. C. (1997). Established and emerging evaluation paradigms for instructional design. In C. R. Dills & A. J. Romiszowski (Eds.), *Instructional development paradigms* (pp. 163-178), Englewood Cliffs, NJ: Educational Technology Publications.

Reinalda, B. (Ed.). (2013). *Routledge handbook of international organization*. New York, NY: Routledge.

Reynolds, M. (2015). (Breaking) The Iron Triangle of Evaluation. *IDS Bulletin*, 46(1), 71-86.

Rogers, E., & Storey, D. (1987). Communication Campaigns. In C. Berger & H. Chaffee (Eds.), *Reader in Public Opinion and Mass Communication* (3rd ed.) (pp. 419-445). New York, NY: Free Press.

Rogers, P. J. (2008). Using programme theory for complex and complicated programs. *Evaluation*, 14(1): 29-48.

Rossi, P., Lipsey, M., & Freeman, H. (2007). Evaluation: A Systematic Approach. (3rd ed.) Beverly Hills, CA: Sage Publications.

Russ-Eft, D., Atwood, R., & Egherman, T. (2002). Use and non-use of evaluation results: Case study of environmental influences in the private sector. *American Journal of Evaluation*, 23(1), 19-31.

Saaty, T. L. (1999). *Decision making for leaders: the analytic hierarchy process for decisions in a complex world* (Vol. 2). Pittsburgh, PA: RWS publications.

Salmon, T. (1989). (Ed.). *Information campaigns: balancing social values and social chang*e. Newbury Park, CA: Sage Publications.

Salmon, T., & Murray Johnson, L. (2001). Communication Campaign Effectiveness. In R. Rice & C. Atkins (Eds.), *Public communication campaigns. 3rd Edition* (pp. 168-181). Newbury Park, CA: Sage Publications.

Salwen, M. B., & Stacks, D. W. (1996). *An Integrated approach to communication theory and research*. Hillsdale, NJ: LEA.

Sanderson, I. (2000). Evaluation in complex policy systems. *Evaluation*, 6(4), 433-454.

Scholes, E. (1997). *Gower handbook of internal communication*. London: Gower Publishing.

Schwandt, T. A. (2003). Back to the rough ground: Beyond theory to practice in evaluation. *Evaluation*, 9 (1), 353–364.

Schwandt, T. A. (2009). Globalizing influences on the western evaluation imaginary. In K.E. Ryan & J.B. Cousins (Eds.), *Sage international handbook on education evaluation* (pp. 119–39). Newbury Park, CA: Sage Publications.

Schwarz, A., & Fritsch, A. (2014). Communicating on Behalf of Global Civil Society: Management and Coordination of Public Relations in International Nongovernmental Organizations. *Journal of Public Relations Research*, 26(2), 161-183.

Scott, R. W. (1995). *Institutions and Organizations*. Newbury Park, CA: Sage Publications.

Scriven, M. (1986). New frontiers of evaluation. *Evaluation Practice*, 7, 7-44.

Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). Newbury Park, CA: Sage Publications.

Shadish, W. R., Cook, T. D., & Leviton, L. D. (1991). *Foundations of program evaluation*. Newbury Park, CA: Sage Publications.

Shadish, W., Newman, D, Scheirer, M., & Wye, C. (Eds.). (1995). *Guiding Principles for Evaluators*. San Francisco, CA: Jossey-Bass.

Shulha, L. M., & Cousins, J. B. (1997). Evaluation use: theory, research, and practice since 1986. *American Journal of Evaluation*, 18(3), 195-208.

Simmons, P. (1998). Learning to Live with NGOs. *Foreign Policy*, *112*, 82-96.

Sireau, N. (2009). *Make Poverty History: Political Communication in Action*. Hampshire, UK: Palgrave Macmillan.

Sriramesh, K. (2009). Globalisation and public relations: The past, present, and the future. *Prism,* 6(2). Retrieved from:
http://www.prismjournal.org/fileadmin/Praxis/Files/globalPR/SRIRAMESH.pdf

Sriramesh, K., & Verčič, D. (2007). The impact of globalization on Public Relations, *Public Relations Review*, *33*(4), 355-359.

Stacks, D. (2006). *Dictionary of Public Relations Measurement and Research*. Institute for Public Relations. Retrieved from: http://www.instituteforpr.org/topics/dictionary-measurement-research/

Stacks, D., & Michaelson. D. (2010). *A practitioner's guide to public relations research, measurement and evaluation.* New York, NY: Business Expert Press.

Stake, R.E. (1980). *Program Evaluation, Particularly Responsive Evaluation. Rethinking Educational Research* (pp.72-87). London: Hodder & Stoughton.

Starling, S. (2010). Monitoring and evaluating advocacy: lessons from Oxfam GB's Climate Change campaign, *Development in Practice*, *20* (2), 277-286.

Stern, E. (2006). Contextual challenges for evaluation practices. In I. Shaw, J. Greene & M. Mark (Eds.), *SAGE Handbook of Evaluation: Policies, Programs and Practices* (pp. 292-315)*,* London: Sage Publications.

Stroup, S. S., & Murdie, A. (2012). There's no place like home: Explaining international NGO advocacy. *The Review of International Organizations*, *7*(4), 425-448.

Sutton, R. (1999). *The policy process: an overview.* London: Overseas Development Institute.

Synder, L. (2001). How effective are mediated Health Campaigns? In Rice, R., & Atkin, C. (Eds.). *Public Communication Campaigns* (pp. 181-192). London: Sage Publications.

Synder, L. (2007). Health Communication Campaigns and Their Impact on Behavior, *Journal of Nutrition Education and Behavior,* Volume 39, Number 2S, March/April 2007, 32-40.

Tkalac, A., & Pavicic, J. (2009). Nongovernmental organizations and international public relations. In K. Sriramesh & D. Verčič (Eds.), *The global public relations handbook: Theory, research, and practice.* 2nd ed (pp. 807-821) New York, NY: Routledge.

Torres, R. T., Preskill, H. S., & Piontek, M. E. (1996). *Evaluation strategies for communicating and reporting: Enhancing learning in organizations*. Thousand Oaks, CA: Sage.

Toth, E., Serini, A., Wright, D., & Emig, A. (2008). Trends in public relations roles: 1990–1995. *Public Relations Review,* Vol. 24, no. 2: 145-163.

Tourmen, C. (2009). Evaluators' decision making. *American Journal of Evaluation,* 30(1), 7-30.

Trochim, W.M.K. (2009). Evaluation policy and evaluation practice. In W.M.K. Trochim, M.M. Mark & L.J. Cooksy (Eds.), *Evaluation policy and evaluation practice*. *New Directions for Evaluation, 123,* 13–32.

Union of International Associations. (2014). *Yearbook of International Organizations 2014-2015, Volume 1*. Boston, MA: Brill Academic Publishers.

Valente, T. (2001). Evaluating communication campaigns. In R. Rice & C. Atkins (Eds.), *Public communication campaigns* (pp. 105-124). Newbury Park, CA: Sage Publications.

Vandenbosch, B., & Higgins, C.A., (1995). Executive support systems and learning: a model and empirical test. *Journal of Managements Information Systems,* 12, 99-130.

Vandenbosch, B., & Higgins, C.A. (1996). Information acquisition and mental models: an investigation into the relation between behaviour and learning. *Information Systems research* 7:198-214.

van Ruler, B. (2015). Agile public relations planning: The Reflective Communication Scrum. *Public Relations Review*, 41(2), 187-194.

van Ruler, B., Tkalac-Verčič, A., & Verčič, D. (Eds.). (2008). *Public relations metrics.* New York, NY: Routledge.

Vogel, I. (2012). *Review of the use of 'Theory of Change' in international development: Review Report.* UK Department for International Development (DFID). Retrieved from: http://r4d.dfid.gov.uk/pdf/outputs/mis_spc/DFID_ToC_Review_VogelV7.pdf

Voils, C. I., Sandelowski, M., et al., (2008). Making Sense of Qualitative and Quantitative Findings in Mixed Research Synthesis Studies. *Field Methods, 20* (1), 3-25.

Walker, G. (1994). Communicating public relations research. *Journal of Public Relations Research,* 6(3), 141–61.

Waltz, K. (1979). *Theory of International Politics.* New York, NY: McGraw Hill.

Watson, T. (1995). *Evaluating public relations: Models of measurement for public relations practice.* Article presented to the International Public Relations Research Symposium, Lake Bled, Slovenia.

Watson, T. (1997). Measuring the success rate: Evaluating the PR process and PR programmes. In P. J. Kitchen (Ed.), *Public relations: Principles and practices* (pp. 283–99). Boston, IL: International Thomson Business Press.

Watson, T., & Noble, P. (2007). *Evaluating Public Relations: a best practice guide to public relations planning, research & evaluation*, 2nd edition. London: Kogan Page.

Weis, J., & Tschirhart, M. (1994). Public information campaigns as policy instruments. *Journal of Policy Analysis and Management*, *13*(1), 82-119.

Weiss, C.H. (1979). The many meanings of research utilization. *Public Administration Review*, 39, 426-431.

Weiss, C.H. (1981). Measuring the use of evaluation. In Ciarlo, J. A. (Ed.). (1981*). Utilizing evaluation: Concepts and measurement techniques* (pp. 17-31). Beverly Hills, CA: Sage Publications.

Weiss, C.H. (1998). *Evaluation: Methods for Studying Programs and Policies* (2nd ed.). Newbury Park, CA: Sage Publications.

Weiss, C.H., Murphy-Graham, E., & Birkeland, S. (2005). An alternate route to policy influence how evaluations affect DARE. *American Journal of Evaluation*, 26(1), 12-30.

Welch, C. (2001). *NGOs and human rights: promise and performance*. Pennsylvania, PA: University of Pennsylvania Press.

Wellman, B. (2002). Designing the Internet for a Networked Society. *Communications of the ACM*, 45(5), 91-96.

White, H. (2008). Of Probits and Participation: The Use of Mixed Methods in Quantitative Impact Evaluation. *IDS Bulletin*, *39*(1), 98-109

White, J. (2005). *Measurement and Evaluation: Moving the Debate Forward*. London: Chartered Institute of Public Relations.

Wholey, J. (1994). Assessing the feasibility and likely usefulness of evaluation. In J. Wholey, H. Hatry & K. Newcomer (Eds.), *Handbook of Practical Program Evaluation* (pp. 15-39). San Francisco, CA: Jossey-Bass.

Wholey, J., Hatry, H., & Newcomer, K. (2010). (Eds.). *Handbook of Practical Program Evaluation*. (3rd ed.). San Francisco, CA: Jossey-Bass.

Willetts, P. (2010). *Non-governmental organizations in world politics: the construction of global governance*. London: Routledge.

Williams, B., & Hummelbrunner, R. (2010). *Systems concepts in action: a practitioner's toolkit*. Stanford, CA: Stanford University Press.

Williams, B., & Imam, I. (Eds.) (2007). *Systems Concepts in Evaluation: An Expert Anthology*. Point Reyes, CA: Edge Press/American Evaluation Society.

Wilson, L., & Ogden, J. (2008). *Strategic Communications Planning for Effective Public Relations and Marketing (5th Edition)*. New York, NY: Kendall/Hunt Publishing Company.

Woolcock, M. (2009). Toward a plurality of methods in project evaluation: a contextualised approach to understanding impact trajectories and efficacy. *Journal of Development Effectiveness*. Vol. 1(1), 1-14.

Xavier, R., Mehta, A., & Gregory, A. (2006). Evaluation in Use: The Practitioner View of Effective Evaluation, *Prism 4*(2). Retrieved from:
http://www.prismjournal.org/fileadmin/Praxis/Files/Journal_Files/Evaluation_Issue/XAVIER_ET_AL_ARTICLE.pdf

Xavier, R., Patel, A., Johnston, K., Watson, T., & Simmons, P. (2005). Using evaluation techniques and performance claims to demonstrate public relations impact: An Australian perspective. *Public Relations Review, 31*(3), 417–24.

Zerfass, A. (2008). A framework for managing and evaluating communication strategies: The Corporate Communications Scorecard. In B. van Ruler, A. Tkalac-Verčič & D. Verčič (Eds.), *Public relations metrics* (pp. 139-153). New York, NY: Routledge.

Zerfass, A., Verčič, D.,Verhoeven, P., Moreno, A., & Tench, R. (2012). *European Communication Monitor 2012. Challenges and Competencies for Strategic Communication. Results of an Empirical Survey in 42 Countries*. Brussels: EACD/EUPRERA.

Zhao, S., Grasmuck, S., & Martin, J. (2008). Identity construction on Facebook: digital empowerment in anchored relationships. *Computers in Human Behavior,* 2008; 24:1816–36.