Guest editorial

1. New avenues in communication evaluation and measurement (E&M): towards a research agenda for the 2020s

1.1 Introduction

The question of how public relations and strategic communication contribute to organisations' success has been a perennial research topic since the 1970s (Stacks, 2017; Gregory and Watson, 2008; Grunig, 2006). Specifically, this question is approached in the sub-field of communication E&M. For more than five decades, scholars have explored how communication effects can be measured alongside different stages (ranging, e.g., from inputs and outputs to outcomes and impacts) and how the value of communication added to the organisational level can be evaluated. According to a recent synthesis of 40 years of literature (Volk, 2016), E&M research has explored numerous questions that can be broadly clustered into the following overarching themes: how to analyse the effectiveness of communication and messages on the output level; how to measure outcomes such as relationships and reputation; how to conceptualise intangible values or capitals and assess the value creation through communication on the impact level; how to further develop E&M methods; and how to assess the state of E&M practices.

As an applied research field, E&M scholars have maintained regular and close relations to the communication industry. Hence, there was and still is a lively and unique exchange between academia and practice (cf. Buhmann *et al.*, 2018). Pioneering scholars, some of them former professionals (earlier on, e.g., Walter Lindenmann, Don Stacks, David Michaelson, Donald Wright or Tom Watson, more recently, e.g., Jim Macnamara or Ansgar Zerfass), have stimulated an vital knowledge transfer with international and national membership organisation (e.g. the International Association for Measurement and Evaluation of Communication (AMEC)), E&M consultancies and client organisations, and various task forces or commissions (e.g. the Institute for Public Relations Measurement Commission, or the International Task Force on Standardisation of Communication Planning/Objective Setting and Evaluation/Measurement Models). Particularly, the last decade saw a considerable amount of collaborative projects, searching for more standardised models, metrics, and methods to advance E&M practices (e.g. recently, AMEC Integrated Evaluation Framework, Barcelona Principles 2.0, Social Media Measurement Standards Conclave, DPRG ICV Framework). However, according to a wide range of studies conducted worldwide to explore the state of E&M in the communication practice, the implementation of such E&M methods and standard metrics focussing on communication outcomes and impact remains low (Baskin *et al.*, 2010; Cacciatore *et al.*, 2016; Macnamara and Zerfass, 2017; Michaelson and Stacks, 2011; Zerfass *et al.*, 2017).

Against this background, E&M scholars have lamented a "stasis" in evaluation practice (Gregory and Watson, 2008) or described the practice as caught in a "deadlock" (Macnamara, 2015). The consequent search for causes has centred on a discussion of possible barriers preventing practitioners from conducting (more sophisticated) E&M, including a lack of: time, budget and resources, competencies and knowledge, management demand and support; access to sophisticated methodologies or tools; and common industry models and standards (cf. Buhmann *et al.*, 2018; Swenson *et al.*, 2019).

Despite the considerable progress of E&M scholarship over the past 50 years, in this Special Issue we postulate the assumption that E&M scholarship can profit from connecting to new developments in the quickly changing communication environment as well as to research conducted in neighbouring fields. We furthermore posit that recent E&M research

can profit from new ideas and perspectives, for instance, when it comes to further exploring barriers to E&M implementation in practice – as most of the extant studies rely on purely descriptive and mostly quantitative data from self-disclosure surveys among practitioners: while these studies help, to a degree, paint a picture of the state of E&M and identify potential barriers, they are susceptible to the self-rationalisations of professionals and do not show empirically how strongly alleged barriers actually affect practitioners' E&M behaviour.

Below, we will first look back and draw attention to current gaps in E&M scholarship. We will then outline the need for new perspectives and for this Special Issue. Following, we will introduce the six contributions to this Special Issue and argue how each of them provides novel ideas for the E&M debate. Finally, looking ahead, we suggest a research agenda for the 2020s in order to inspire future E&M scholarship to shed light on some of the unresolved questions, break new ground and draw insights from related disciplinary perspectives.

2. The current state of E&M research

While practices of public relations measurement can be traced back to the late eighteenth century, scholarly research and theorisation of E&M began to evolve in the late 1960s and made substantial progress through the 1970s and 1980s (for a historical overviews, see e.g. Likely and Watson, 2013; Watson, 2012). Over the course of 50 years, questions related to E&M have neither lost momentum nor relevance for scholars. Earlier research, until the 2000s, focussed mainly on analyses of communication effectiveness, methodological refinements, the development of E&M models and frameworks for practice, and empirical studies of evaluation practices. While these topics continue to be prominent in the literature until today, the past two decades were marked by a trend towards importing insights from management and organisation literature, and a stronger orientation towards challenges in measuring outcomes (e.g. relationships or attitude change) as well as impact at the organisational level (e.g. reputation or brand value).

2.1 Current issues and gaps

In view of the long history of communication E&M research, recent works have summarised the debate and highlighted prevalent issues and gaps in the literature (cf. Buhmann and Likely, 2018; Volk, 2016; Likely and Watson, 2013; Macnamara, 2014). While some of these issues are geared more at the state of the practice (such as lack of outcome measurement), others focus more on the state of E&M research. From the latter we can distil the four following overarching gaps in the literature.

Heterogeneous terminologies and non-standardisation in E&M. PR and communication researchers have been closely involved in the development of various frameworks for the practice that represent the (ideal) evaluation process, such as Cutlip *et al.*'s (1985) often-cited Planning, Implementation, Impact Model, Lindenmann's (1997) PR Effectiveness Yardstick or Watson's (1997) Short Term Model and Continuing Model of Evaluation to name just a few of the early models. This has been done by employing a plethora of concepts and literature for distinguishing main phases, stages and units of evaluation. Yet, in spite of recent efforts aimed at synthesis, definitional inconsistencies remain and key constructs and terms – even "evaluation" and "measurement" – are used synonymously or confusingly (Macnamara and Likely, 2017; Schriner *et al.*, 2017).

The prevalent heterogeneity of available frameworks and measures has led to a growing critique of a lack of standards (Macnamara, 2014; Michaelson and Stacks, 2011; Ragas and Laskin, 2014), also voiced by leading practitioners (Marklein and Paine, 2012). Countering the criticism, scholars have recently called for proposing more integrated and holistic models as well as developing unified "inventories" of measures and metrics (cf. Macnamara

and Likely, 2017; Schriner *et al.*, 2017). Nevertheless, little is known to date about how standardisation processes evolve and what actually makes standards stick (or not).

Overemphasis on the instrumental focus. Most research in E&M follows a functionalist, positivist and partially normative approach (Volk, 2016; Macnamara, 2014), which can be attested, e.g., by quantitative practitioner surveys assessing "the state of the field" or the development of applicable best practice models for E&M. The prevalence of rationalistic and instrumental assumptions has recently been problematised for being too narrow, control focussed and organisation centric, combined with a call for more open, continuous, dynamic and expanded approaches (Macnamara and Gregory, 2018). Yet, critical inquiries and qualitative and interpretive approaches remain strongly underrepresented in the E&M literature (Macnamara, 2014).

Moving beyond instrumental towards critical questions could provide novel insights for the evaluation body of knowledge. However, few works have so far explored the possible "pathologies" in specific evaluation practices and of evaluation as a whole. For instance, previous research has shown that communicators may feign expertise in evaluation, whitewash data and produce invalid or unreliable reports (Place, 2015). From this perspective, evaluation becomes rather a form of self-justification that fixates on performance, rationality and objectivity to stabilise some particular ideology. It is conceivable that evaluators shade, over generalise, or tamper results to support particular goals or decisions in compliance within an established power structure. Such "pseudoevaluations" are a recognised issue in evaluation in many other domains (Stufflebeam and Coryn, 2014), but are rarely addressed in PR and communication E&M literature.

Disregard of intervening variables and prevented/hidden/indirect effects. Alongside the dominant orientation towards research with application value for practice comes the risk of overly simplistic, mechanistic and linear assumptions of communication effects. Many extant E&M models and frameworks ignore intervening variables, do not rigorously separate antecedents from outcomes of communication processes and postulate unwarranted direct cause–effect relationships. The perspective of the recipient and the social and organisational context, in which communication messages and effects are embedded, are largely overlooked (Buhmann and Likely, 2018). For instance, the mainstream E&M literature has focussed on intended messaging effects, but has remained widely silent on elusive hidden, indirect or prevented effects of communication. This is quite surprising, since preventing communication – i.e. a critical media coverage – is a legitimate and often-pursued goal of communication professionals. Following Wehmeier and Nothhaft (2013, pp. 123-124), one may even claim that PR E&M scholars may run a two-sided risk in developing E&M models and frameworks that oversimplify regarding third and intervening variables: first, such models may fail to reliably measure communication effects in practice, and hence will not aid practitioners, possibly decreasing their confidence in academic knowledge. Second, E&M scholarship risks to be perceived as an industry-driven, atheoretical field, downgrading its standing in the wider domain of PR scholarship.

Siloed debate and disconnect from other disciplines. Further criticism has evolved around the siloed approaches to E&M theorising and research. Even though the past two decades were marked by a growing import of managerial thinking (e.g. the resource-based theory of the firm or stakeholder theory, see Volk, 2016), the contemporary debate remains quite fragmented and unconnected to recent advancements in its parent discipline – communication science – as well as its embedding in the organisation context – studied in management, organisation and marketing/advertising research. One recent attempt to reach out for insights from one of the neighbouring fields has been put forward by Macnamara and Likely (2017), who suggested a disciplinary "home visit" to programme

evaluation that could inform the search for standards and overcome the long-standing stasis in practice. Strikingly, the majority of E&M literature has neglected to import highly relevant knowledge generated in the neighbouring disciplines and communication subfields.

2.2 How far have we come?

Going beyond a list of current issues and gaps in E&M research, and attempting to further assess how far E&M research has come we draw on Nothhaft *et al.* (2018), who refer to Shneider's (2009) four-stage model for describing the evolutionary life cycle progress of scientific fields. Shneider (2009) suggests four stages to model the typical life cycle: in brief, Stage 1 describes the invention of a discipline and the development of a new scientific language and respective theoretical frameworks. In Stage 2, inventive scholars build on the insights of the first-stagers and develop them further to achieve breakthroughs in terms of methods and techniques. Stage 3 constitutes the apex of a discipline where methods and theories begin to inform other disciplines, standards are developed and tolerance for mistakes is low. In Stage 4, disciplines are reaching maturity and aim at safeguarding the body of knowledge in order to claim a state of the art, with a focus and refining pedagogy (Shneider, 2009, pp. 218-221; Nothhaft *et al.*, 2018, pp. 354-355).

Of course, transferring this four-stage model has to be made with caution, not only because E&M research is rather a sub-field of PR and strategic communication scholarship than a discipline in its own right, but also because the disciplinary progress and status of the field as a whole is indeed debatable (see e.g. Nothhaft *et al.* 2018). According to Nothhaft *et al.* (2018), however, the E&M debate can be characterised as a model example of Stage 2, in which "energetic contributors […] not only develop the new language further, but develop methods and techniques that yield concrete, increasingly precise results" (p. 355). They argue that the E&M field comprises a "rather small and tightly knit" (p. 356) community, in which "colleagues work in cooperation with the industry to develop systems that are not only academically sound but work in practice" (p. 355), but they also criticise the "unsatisfactory transfer between academia and practice and practice more inspired by academia rather than academics taking the lead" (p. 356).

Indeed, the previous review of the achievements and shortcoming in the E&M field is in line with Nothhaft *et al*'s assessment. We see a fairly well-defined but small community of scholars and professionals (Likely, 2018), who work on standardised models and metrics and a unified terminology – which would be one characteristic of a Stage 3 discipline. Nevertheless, E&M scholarship has remained a niche research field within the broader PR/strategic communication field – especially when compared to the popular subfields of crisis communication or OPR and dialogic theory – and consequently, a marginal side topic in the broader communication discipline. A commonly agreed theoretical basis is still missing (Likely and Watson, 2013; Volk, 2016), although there are constant conceptual developments. Little knowledge has so far been imported from neighbouring disciplines and E&M research is still far from exporting knowledge or methods and informing other disciplines. Hence, E&M research has clearly not yet reached Stage 3; however, considering the range of popular textbooks and handbooks available, there are certainly attempts to safeguard the produced knowledge, which is characteristic for Stage 4. Looking ahead, in order to progress towards Stage 3, new conceptual approaches, refined methodologies and the inclusion of knowledge produced in related disciplines are necessary for the E&M community.

3. Rationale of the Special Issue

This Special Issue emerged from a panel at the International Communication Association's (ICA) 2018 Annual Conference in Prague, which was organised by the first author, titled "New voices in PR evaluation: innovative approaches and new research avenues for a field in stasis". The rationale of this panel was that "new voices" are needed to move E&M out of

the diagnosed "deadlock" or "stasis" and to stimulate future progress in scholarship. The panellists – some of them contributors to this Special Issue – presented provocative ideas and new perspectives mostly inspired by work outside of the core PR research domain.

The underlying assumption was that, in spite of considerable scholarly efforts, the broad majority of E&M research has explored a fairly focussed set of research questions with and largely sustains from challenging existing explanations or developing alternative assumptions. In following Alvesson and Sandberg's (2014) discussion of "boxed-in research", we posit that much E&M research has been conducted "in the box". While the positive development of the E&M community towards sharing a common research interest and coherent body of knowledge is generally appreciated, signalling the further maturation of this sub-discipline, we share concerns voiced by Alvesson and Sandberg regarding the limitations of boxed-in research (see also Werder *et al.*, 2018, pp. 335-336). Despite the many advantages of boxed-in research, Alvesson and Sandberg (2014, p. 974) raise their concerns against a possible shortage of imaginative and interesting research, overspecialisation, silo mentality, fragmentation, box identity, suspiciousness, intra-box communication, unnecessary polarisation among researchers and unquestioning attitudes. Research conducted within a specific box is "characterized by a strong and narrow focus on some issues within a well-defined, specialized intellectual terrain" (Alvesson and Sandberg, 2014, p. 971). Some of these negative consequences of boxed-in research may apply to E&M scholarship just as well – for instance, the occurrence of a silo mentality and fragmentation, or the shortage of interesting research that questions dominant concepts or attitudes. Indeed, this is the key argument put forward by Nothhaft and Stensson (2019) in this Special Issue, who criticise the "circumspection and narrowness" of the debate about barriers to E&M implementation, and claim that the academic debate shows symptoms of "functional stupidity".

In view of this assessment, we see a need to support more box-changing research that has the potential to broaden the debate. Box breaking research occurs when scholars "broaden their intellectual territory and research competence by moving beyond their specific research boxes to also consider the resources and ideas of other research boxes and intellectual terrains" (Alvesson and Sandberg, 2014, p. 968). This exactly is the ambition for this Special Issue: to provide a platform for unconventional and alternative arguments, interdisciplinary and integrative perspectives. We therefore particularly invited submissions that would introduce new arguments and highlight promising starting points for a more "diverse" future of E&M scholarship.

4. Articles in the Special Issue
Against the backdrop of the above discussion on the state of communication E&M, we turn to the six articles contained in this Special Issue. In our view each article makes important contributions to the current advancement of the debate, opening up new and interesting angles for understanding the practices and contexts of E&M.

Table I provides a summary of each contribution's primary purpose, theme, key literature, approach and implications. In contrast to the overall trend towards a fixation on (quantitative) empirical research in communication science as well as in strategic communication/public relations, most of the invited articles in this volume set a counterpoint and present conceptual arguments, partially supported by qualitative insights. In terms of disciplinary orientation and key literature, most contributors drew considerable inspiration from organisation research and management science. With regard to the themes, four of the six articles shed light on the application of E&M in communication practice in general or in specific organisations, examining different barriers that might hinder the implementation or advancement of E&M, i.e. organisational, motivational, cultural, structural, technical or alignment-related barriers. These articles have significant overlaps at the outset, but propose quite alternative explanations for the existence of barriers, and

Article	Theme	Primary purpose	Key literature/theory	Method/Approach	Implications
Kim and Cappella: "Reliable, valid and efficient evaluation of media messages: developing a message testing protocol"	Message effectiveness, methodology	To propose a reliable, valid and efficient standard process for the evaluation of media messages, based on recent advances in message effects research and campaign design research	Message effects, message testing, persuasion	Meta-analysis	Use perceived message effectiveness (PME) and perceived argument strength (PAS) as strong proxies allowing for inferences about relative effectiveness when many messages need to be evaluated quickly and efficiently
Sommerfeldt and Buhmann: "The Status quo of evaluation in public diplomacy: insights from the US State Department"	E&M practices	To understand the barriers and context around the reluctance in the practice to evaluate public diplomacy programmes	Public diplomacy concepts	Qualitative interviews	Assess cases of reluctant application of M&E by addressing potentially disjointed systems and structures, lack of development and focus, devaluation of public diplomacy efforts, lack of resources, capital-field tensions, reliance on outputs and anecdotes, as well as misalignment of public diplomacy goals
Nothhaft and Stensson: "Explaining the measurement and evaluation stasis: a thought experiment and a note on functional stupidity"	E&M practices	To explain the E&M stasis in practice through consideration of practitioners' self-interest as business people and the industry's anomaly of expectation-inflation and overpromising	Principal–agent theory, theory of lemon markets, functional stupidity	Thought experiment, qualitative interviews	Test the plausibility of motivational barriers at the individual level as an explanatory factor for the stasis of E&M practices; avoid narrow and circumspective research assumptions and functionally stupid explanations in theorising about barriers in practice
Murtarelli, Romenti, Miglietta, and Gregory: "Investigating the role of contextual factors in effectively executing communication evaluation and measurement: a scoping review"	E&M practices	To conceptualise the role of internal contextual factors affecting the E&M management process as an often-overlooked barrier of E&M practices	Programme evaluation theory, theory-based evaluation, theory of change	Narrative scoping review	Consider the so far often neglected contextual factors influencing the practice of E&M: (a) evaluative capacity and history (organisational context); (b) evaluative culture and leadership (cultural context); (c) stakeholder–evaluator relationship (relational context); (d) evaluation communicative network (communicative context)
Gilkerson, Swenson and Likely: "Maturity as a way forward for improving	E&M practices	To explicate the concept of maturity and propose a maturity model that	Maturity models and concepts	Conceptual review	Use the dimensions of a "holistic approach" (i.e. levels of analysis), "investment" (time, budget, tools, skills), "alignment" (process, integration) and "culture"

(continued)

Table I.
Overview of articles in this Special Issue

Article	Theme	Primary purpose	Key literature/ theory	Method/ Approach	Implications
organisations' communication evaluation and measurement practices: a definition and concept explication"		highlights the key dimensions of E&M maturity in organisations and can serve as a basis for maturity assessments			(motivations, relationships, standards) to assess E&M maturity in organisations as a step towards, increasing both accountability and credibility for the work done by the communications function
Van Ruler: "Agile communication evaluation and measurement"	Methodology, E&M practices	To question the fundamental assumptions of summative, goal-dependent and organisation-centric evaluation in view of agility concepts, and to present alternative measurement methods addressing the needs of agile organisations	Agility concepts, evaluation theory, action research, sense-making methodology, Weberian idea of "Verstehen"	Development debate	Reconsider E&M theory and practice against the radical assumptions of agility, moving towards goal-free and on-going evaluations oriented towards users/needs and qualitative measurement methods, such as action research and sense-making methodology

Table I.

outline different implications for how our theorising on barriers should develop further. The other two articles focus specifically on methodological approaches to conducting E&M. While one article investigates the basic mechanisms of message effects, going back to the roots of communication science, the other article challenges the fundamental assumptions of E&M scholarship in view of the changing organisational and communicative environment.

4.1 Annotation of the articles in this Special Issue
The first study revives an older, largely abandoned debate in the E&M literature. Evaluation of message effectiveness has been one of the initial starting points of E&M research, at least since the 1970s. In fact, message effectiveness was among the most researched topics in the early 1980s and 1990s (Volk, 2016). However, the topic received little attention over the past decades and the field has not kept abreast to the discussion on message testing as it evolved especially in message effect, campaign design and persuasion research. Kim and Cappella (2019) introduce and discuss an efficient, reliable and valid message testing protocol covering how to conceptualise and evaluate the content and format of messages; procedures for acquiring and testing messages; and the application of robust measures of perceived message effectiveness and perceived argument strength. The particular strength of the approach lies in the ability to facilitate a selection of candidate messages for subsequent deeper testing, for various types of communication campaigns and for research in theory-testing contexts – avoiding the limitations of using a single instance of a message to represent a category (i.e. case-category confound). Ultimately, the contribution also shows that strong inspiration that can be drawn from adjacent fields, such as health communication, that have developed very sophisticated approaches with high potential for adaptation within the domain of communication E&M.

Sommerfeldt and Buhmann (2019) posit that, despite increased calls for enhanced monitoring and evaluation in public diplomacy, the state of practice remains grim. To better understand this situation, they undertook a qualitative study uncovering the perceptions around evaluation from the voices of those who must practice it. The authors conducted 25 interviews with public diplomacy officers with the US Department of State, both in Washington, DC as well as at posts around the world. The study shows a previously not discussed tension between diplomacy practitioners in Washington, DC and those in the field stationed outside the USA. Specifically, this pertains to a lack of joined clarity about the goals of specific public diplomacy programmes and public diplomacy as a function in general – which form the basis of targeted evaluations. The practice-level insights from this study contribute to a better understanding of the factors that may drive or hamper evaluation in day-to-day public diplomacy work. Ultimately, the article also shows that, while communication E&M face some persistent challenges which reach across professional fields (such as lack of resources, reliance on outputs or misalignment of goals), there are also unique challenges (such as capital-field tensions) that pertain to field-specific contexts and that need to be understood more closely to a particular domain of communication practice.

Nothhaft and Stensson (2019) propose a new angle to explaining the evaluation "deadlock" or "stasis" and the discrepancy between practitioners' words and actions when it comes to E&M. While, at present, most explanations tend to focus on lack of knowledge or inadequate systems or frameworks, the authors – by means of a thought experiment and qualitative interviews with practitioners – highlight practitioners' self-interest as well as their passive or active (but covert) forms of resistance against E&M. Using the functional stupidity concept, the authors argue that the current academic debate is conducted in narrow and circumspect ways and that we, as academics, should "take the blinders off" to allow for alternative explanations. That is, that a critical mass of actors benefits from talking about E&M, yet stands to lose from doing it. The authors posit that if the long-time neglect of E&M has led to expectation-inflation and overpromising, even well-performing actors might shy away from

rigorous E&M, fearing being measured against promotional, not realistic standards. At the same time, on the level of industry discourse, these practitioners would still advocate for E&M in principle, so as to avoid the suspicion of underperformance.

Romenti *et al.* (2019) posit that while most work aims to raise the level of sophistication in practice – discussing best practice models and methods – in real life, most organisations experience issues in implementing and managing E&M that remain largely unexplored. Bringing together insights from both the programme evaluation and performance measurement literature, the paper works out key organisational, relational, cultural and communicative factors, which constitute the context of management of the evaluation process and ultimately come to bear on any E&M implementation. The identified contextual factors are clustered into four groups: evaluative capacity and history (organisational context); evaluative culture and leadership (cultural context); stakeholder–evaluator relationship (relational context); and evaluation communicative network (communicative context). Thus, the authors provide new arguments for additional and alternative dimensions in the E&M barriers/drivers literature (e.g. Buhmann and Brønn, 2018; Macnamara, 2015).

Gilkerson *et al.* (2019) argue that a stronger focus on the maturity concept in E&M has the potential to advance the field by increasing both accountability and credibility for the work done by the communications function. While general "maturity stages" (adolescent, advanced and mature) have been used several times to assess the advancement of E&M in organisations, this literature has not defined maturity and its dimensions. Drawing from previous work on maturity models within other fields, recent communication scholarship and industry practice, the authors propose a new theoretical conceptualisation of maturity, including the construct's core dimensions and sub-dimensions. E&M maturity is conceptualised into four essential elements: holistic approach, investment, alignment and culture. The contribution of E&M efforts is represented as the direct support of corporate strategy, and ultimately increased value, from the communications function. Operational elements of maturity include levels of analysis, time, budget, tools, skills, process, integration, motivations, relationships and standards.

Van Ruler (2019) posits that, while most E&M models and approaches rely on the assumption that specific and quantifiable goals precede the development of any practical E&M, the reality of organisational processes can be a lot more fluid and agile, often with fast-moving targets, which calls for an introduction of the concept of agility. Based on a literature review, the author highlights central tenants of agile E&M, such as that what works is more important than what was agreed upon in advance, that agile evaluation is always formative (as opposed to summative), and that qualitative methods (e.g. action research or sense-making approaches) are key to supporting formative and fluid practices. These points challenge widely established (organisation-centric) concepts of E&M and brings the needs of users and the context of a rapidly changing environment into the equation.

4.2 Comment on the box-changing contributions of the articles

In Figure 1, we visualise how the articles in this Special Issue present research that can be characterised as box changing. According to Alvesson and Sandberg (2014), a box-changing work is rooted in a specific box, but "reaches outwards for new ideas, theories or methods that can be used to change the box in some significant way" (p. 980). The aim is hence not to further refine the existing body of knowledge, but instead to "identify crucial problems that can lead to substantial rethinking about central elements in the existing literature". Alvesson and Sandberg categorise this as "a radical reform from within" (p. 980), but acknowledge that box changers still remain boxed-in to a certain extent, observable, e.g. in terms of a similar style, references, vocabulary or community, where scholars attend the same conferences or publish in the same journals. One step forward in this metaphor is box jumping research, which occurs when scholars leave their primary box to work with two or
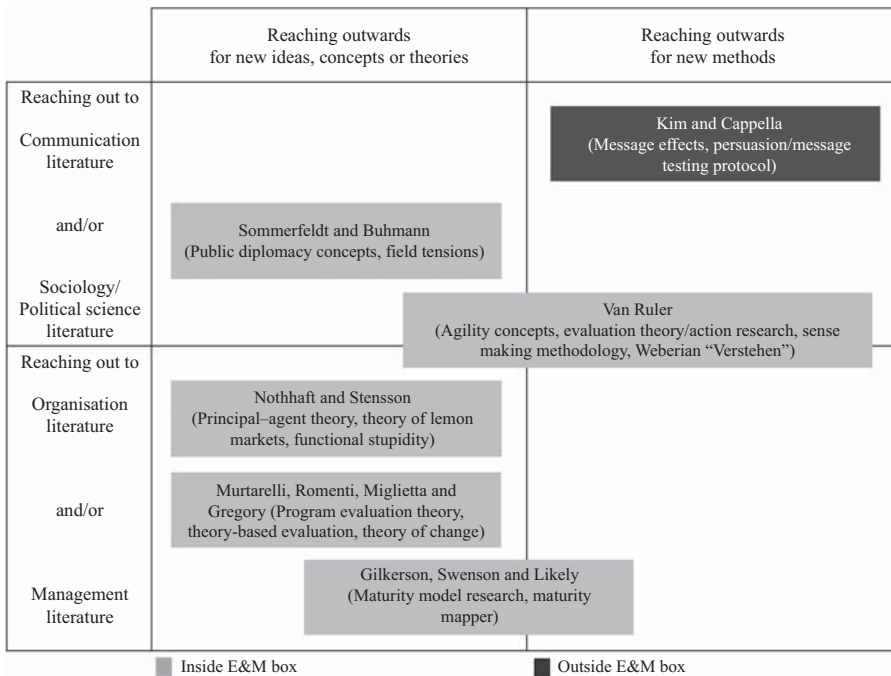
	Reaching outwards for new ideas, concepts or theories	Reaching outwards for new methods
Reaching out to Communication literature		Kim and Cappella (Message effects, persuasion/message testing protocol)
and/or	Sommerfeldt and Buhmann (Public diplomacy concepts, field tensions)	
Sociology/ Political science literature		Van Ruler (Agility concepts, evaluation theory/action research, sense making methodology, Weberian "Verstehen")
Reaching out to Organisation literature	Nothhaft and Stensson (Principal–agent theory, theory of lemon markets, functional stupidity)	
and/or	Murtarelli, Romenti, Miglietta and Gregory (Program evaluation theory, theory-based evaluation, theory of change)	
Management literature	Gilkerson, Swenson and Likely (Maturity model research, maturity mapper)	

■ Inside E&M box   ■ Outside E&M box

Figure 1.
Box-changing articles
in this Special Issue

several boxes and engage with significantly different topics, methods or theories simultaneously (cf. Alvesson and Sandberg, 2014, pp. 980-982). Even a step further, box transcendence, occurs when scholars aim for the "synthesis, confrontation, concept blending or bricolage of various metaphors or empirical materials".

In light of the above, all articles in this Special Issue meet the criteria of box-changing research. They take the E&M literature as a primary reference point and then reach outwards for new contexts, ideas and theories or for new methods and draw inspiration from neighbouring fields. Using these new perspectives, they challenge the E&M box or rework box-specific assumptions. However, they also show characteristics of boxed-in research, since they share the same vocabulary, references and scholarly community. One exception is the work by Kim and Cappella (2019) who come from media effect and health communication research, but bridge disciplinary silos by engaging with the E&M literature.

The four articles focussing on E&M barriers in communication practice reach out for new contexts, ideas and theories and provoke a substantial rethinking of how we theorise the diagnosed stasis. Nothhaft and Stensson (2019) question the dominating assumption that E&M practices are inhibited by a lack of budget, time, knowledge or methods, and instead propose an alternative and provocative assumption that rigorous E&M might simply not be part of practitioners' self-interests. Using a thought experiment they draw on the concepts such as "lemon markets", principal–agent theory, and functional stupidity to emphasise that this explanation for the E&M stasis has been overlooked by scholars so far.

The works by Romenti *et al.* (2019) and Gilkerson *et al.* (2019) identify further crucial problems in the current debate about barriers – the omission of the organisational context in which communication E&M is embedded – and reach out towards programme evaluation, performance measurement literature, and maturity conceptualisations and models for inspiration. Instead of merely adding to the literature on barriers by

conducting another quantitative study diagnosing a stasis in practice, they introduce novel theory-based explications. Their findings imply that future investigations into E&M barriers need to consider further contextual factors and go beyond self-disclosed surveys of communication professionals.

Sommerfeldt and Buhmann (2019) add to the research on barriers from a much-needed qualitative perspective, producing insights form the PR-related field of public diplomacy. Their findings highlight resemblances with studies conducted in the professional fields of strategic communication and public relations more generally, and they point out specifically to so far overlooked tensions regarding E&M between central and peripheral units of organisation.

Van Ruler (2019), in contrast to the common emphasis on the relevance of goal-setting for evaluations (cf. e.g. Hon, 1998 but also Sommerfeldt and Buhmann in this Special Issue), challenges the very fundamental assumptions of "mainstream" E&M research, and criticises dominating summative and organisation-centric approaches to E&M. Van Ruler instead reaches out to agility research and calls for a shift towards action research or sense-making methodologies.

5. A research agenda for the 2020s

The articles collected in this Special Issue open up promising new and alternative avenues for future research and demonstrate the potential for the continuous reflection of already established assumptions in the "E&M research box". Looking ahead, we believe first and foremost that research in the 2020s can profit from increased efforts particularly to understanding and explaining E&M and its manifestations in practice, rather than aiming at improving practices. Of course, producing theories with a "cash value" (Toth, 2002) is a legitimate goal for an applied research field such as E&M and there is indubitably a need in practice for scientifically sound but efficient methods, measures, models, frameworks and tools (Wehmeier and Nothhaft, 2013). Grunig (2006, p. 152) for instance famously claimed "public relations scholars need to develop both positive and normative theories – to understand how public relations is practiced and to improve its practice […]". But improving practices and fostering professionalism through the development of positivist, prescriptive, functionalist or normative scholarship should not be the only goal. Fundamental research into the underlying logics of E&M practices and critical and interpretive examinations are just as relevant – and currently still starkly underemphasised.

Following this line of thought, and responding to earlier calls (e.g. Wehmeier and Nothhaft, 2013; Aggerholm and Asmuß, 2016), we advocate for research employing a practice perspective on E&M theorising and studies (e.g. see Whittington, 2007; Czarniawska, 2008; Jarzabkowski *et al.*, 2007; Sandberg and Tsoukas, 2011). We believe that a turn towards a practice perspective would allow scholars to better reflect the logics and (ir)rationalities of practice and, e.g., shed new light on the often-discussed barriers of E&M practices. By observing and studying how E&M practitioners act and interact in the organisation on a day-to-day basis and with which motivations and consequences, scholars would gain a deeper understanding of the social mechanisms underlying E&M practices (Sandberg and Tsoukas, 2011; Aggerholm and Asmuß, 2016).

Looking ahead, below we outline a research agenda with eight possible directions for future research for the 2020s. This agenda builds on both the existing and established discussion on the state of the field in general as well as on the suggestions proposed in consequence of the 2018 ICA panel discussions, and the work presented in this Special Issue:

(1) Towards clarifying key terminology and understanding standardisation processes:

- How can vague, inconsistent definitions and diverging approaches to operationalisation of key constructs be harmonised? What should be the basic definitional elements of key E&M terms in the field?

- How do standards as formulated rules for common and voluntary uses actually evolve from their early development to later application and compliance (or not)? What is the role of central actors in standard setting and what are dynamics of the process of standardisation in E&M?

- What is the value of scalable (or customised) standards for E&M practice? What is the value of accessible standardized databases in which E&M data collected across organisations could be integrated, synthesised, and compared?

(2) Towards fundamental research aimed at understanding E&M practices:

- How are E&M frameworks/models and methods actually used in the practice (with what motivations, to what end)? How do practitioners get away with avoiding evaluation or using invalid or vanity metrics? (Macnamara, 2018).

- What is the power/pressure context in which E&M practices evolve (or not)? Which variables explain the variability in the adoption of (mature) E&M systems?

- What are the drivers and barriers to conduct E&M, at the industry level (i.e. anomaly of overpromising), organisational level (i.e. evaluation culture, management support, evaluation capacity, history, structures, etc.), department level (i.e. leadership, alignment, stakeholder relations, life cycle, etc.) and individual level (i.e. motivations, expertise, etc.)? (Romenti and Murtarelli, 2018)

- How do E&M fads and fashions evolve and spread (and disappear)? How do challenges of implementing evaluation correspond to other related practical challenges, i.e. the formulation of strategically aligned goals and their operationalisation into measurable objectives?

(3) Towards understanding intervening variables in E&M:

- Which theoretical frameworks and concepts from communication science (i.e. priming, cultivation, agenda setting, information processing, persuasion, etc.) can be integrated in E&M frameworks and models to better conceptualise the nature of communication processes?

- Which studies can inform the modelling of relations between antecedents and outcomes of communication in order to make more cautious interpretations of causality?

- How can we better explain the impact of mediating factors, including, e.g. psychological factors (e.g. influence of personality traits on perceptions and emotions, cognitive dissonance) or social factors (e.g. influence of social networks and group integration on attitudes)?

(4) Towards understanding prevented or hidden communication effects:

- Which frameworks and models (i.e. crisis/risk communication, conflict resolution, etc.) provide a basis for conceptualising and measuring prevented communication effects? How do hidden or indirect communication effects manifest empirically?

- How can internal evaluation services (i.e. media intelligence, strategic insight analysis) provided by communication departments to inform organisational decision making be integrated in conceptualisations of E&M frameworks? How can the value of counselling and advising organisational leaders, based on communication E&M insights, be conceptualised?

(5) Going beyond positivist epistemology:

- Which paradoxes or conflicting patterns of E&M have been discussed in the literature and can be observed in practice? In which contexts do communication E&M logics conflict with the overall logic of the organisation in relation to evaluation (i.e. reluctance to quantification vs measurement-driven cultures, etc.)? (Romenti and Murtarelli, 2018)

- How are E&M results used for ritual, legitimacy purposes of signalling rationality (i.e. pseudoevaluations, myths of rationality)? How do E&M tools and results shape and change perceptions of organisational realities (i.e. performativity)? Which uses of E&M results are harmful for the organisation (i.e. overly trust in data, blind spots, functionally stupid evaluations, etc.)? (e.g. see Falkheimer *et al.*, 2016; Wehmeier, 2006)

(6) Incorporating critical thinking:

- Which perspectives (i.e. critical, sociocultural, postmodern, etc.) can help to overcome the extant functionalist organisation-centric assumptions? What are the shortcomings and deficiencies of contemporary theorising about the fundamentals of measurement of communication effects (e.g. Thummes, 2009)? How can we better conceptualise the (unintended) societal implications of communication effects?

- Which ethical quandaries exist in E&M practices (i.e. misrepresentation or whitewashing of data, deceiving data display, use of vanity metrics, e.g. Place, 2015)? In which cultural situations or industry contexts do unethical E&M practices evolve? What are the hidden motivations of practitioners and which consequences arise from unethical E&M behaviour?

(7) Exploring and criticising existing measures and metrics:

- Which metrics assist in the measurement and valuation of immaterial capitals (i.e. social capital, human capital and intellectual capital) (e.g. see De Beer, 2014)?

- Which measurement methods and metrics are applicable and useful in agile organisational contexts with dynamic communication goals and fluid organisational targets?

- How can the cognitive and behavioural measures typically used in communication E&M be made compatible with overarching organisational business, performance and valuation metrics? How can advancements in other professional fields (i.e. online marketing/advertising, data science and artificial intelligence, accounting and controlling standards, etc.) inform the further development of metrics for communication E&M (i.e. social media metrics, big data analytics, etc.)?

(8) Bridging disciplinary siloes and aiming for synthesis:

- Which recent discussions in the neighbouring disciplines can be integrated in our research field to resolve the continuing fragmentation into siloed debates and make our research more compatible, particularly from…?

  – media effect research, media psychology, public opinion/audience/reception research, health communication and public diplomacy;

  – programme evaluation and performance measurement/management;

– critical organisation and management research;

– strategy and alignment research;

– (online/social) marketing measurement, market research, brand evaluation and advertising effectiveness research;

– accounting and controlling research; and

– data science and data analytics.

- And vice versa, how can we enrich and inform the evaluation debates carried out in neighbouring fields with findings from our field (i.e. by re-conceptualising simplistic understandings of communication processes)?

- Which economic theories help to better reflect the economic dimension of the organisational setting in which communication effects and E&M takes place?

- Which implications will blurring boundaries and increasing convergence among communication fields, i.e. marketing, branding, advertising, journalism and strategic communication, have for E&M?

Hopefully, this collection of eight directions for future research as well as the arguments provided in the six articles in this Special Issue might inspire scholars to walk some new avenues in the 2020s. We hope that further work will follow the signposts for box-changing research and consider the resources and ideas of other disciplines' research boxes and intellectual terrains, thus potentially contributing to further opening up the E&M research box.

Sophia Charlotte Volk
Institute of Communication and Media Studies, University of Leipzig,
Leipzig, Germany, and
Alexander Buhmann
BI Norwegian Business School, Oslo, Norway

References

Aggerholm, H.K. and Asmuß, B. (2016), "A practice perspective on strategic communication: the discursive legitimisation of managerial decisions", *Journal of Communication Management*, Vol. 20 No. 3, pp. 195-214.

Alvesson, M. and Sandberg, J. (2014), "Habitat and habitus: boxed-in versus box-breaking research", *Organisational Studies*, Vol. 35 No. 7, pp. 967-987.

Baskin, O., Hahn, J., Seaman, S. and Reines, D. (2010), "Perceived effectiveness and implementation of public relations measurement and evaluation tools among European providers and consumers of PR services", *Public Relations Review*, Vol. 36 No. 2, pp. 105-111.

Buhmann, A. and Brønn, P. (2018), "Applying Ajzen's theory of planned behaviour to predict practitioners' intentions to measure and evaluate communication outcomes", *Corporate Communications: An International Journal*, Vol. 2 No. 2, pp. 1356-3289.

Buhmann, A. and Likely, F. (2018), "Evaluation and measurement", in Heath, R.L. and Johansen, W. (Eds), *The International Encyclopedia of Strategic Communication*, Vol. 1, Wiley-Blackwell, Malden, MA, pp. 625-640.

Buhmann, A., Likely, F. and Geddes, D. (2018), "Communication evaluation and measurement: connecting research to practice", *Journal of Communication Management*, Vol. 22 No. 1, pp. 113-119.

Cacciatore, M., Meng, J. and Berger, B. (2016), "Measuring the value of PR? An international investigation of how practitioners view the challenge and solutions", paper presented at 19th International Public Relations Research Conference, Miami, FL, 2–6 March, available at: www.iprrc.org/proceedings (accessed 22 April 2019).

Cutlip, M., Center, A. and Broom, G. (1985), *Effective Public Relations*, 6th ed., Prentice Hall, Englewood Cliffs, NJ.

Czarniawska, B. (2008), *A Theory of Organising*, Edward Elgar Publishing, Cheltenham.

De Beer, E. (2014), "Creating value through communication. Public relations and communication management in South Africa", *Public Relations Review*, Vol. 40 No. 2, pp. 136-143.

Falkheimer, J., Heide, M., Simonsson, C., Zerfass, A. and Verhoeven, P. (2016), "Doing the right things or doing things right?", *Corporate Communications: An International Journal*, Vol. 2 No. 2, pp. 142-159.

Gilkerson, N.D., Swenson, R. and Likely, F. (2019), "Maturity as a way forward for improving organizations' communication evaluation and measurement practices: a definition and concept explication", *Journal of Communication Management*, Vol. 23 No. 3, pp. 246-264.

Gregory, A. and Watson, T. (2008), "Defining the gap between research and practice in public relations programme evaluation – towards a new research agenda", *Journal of Marketing Communications*, Vol. 14 No. 5, pp. 337-350.

Grunig, J.E. (2006), "Furnishing the edifice: ongoing research on public relations as a strategic management function", *Journal of Public Relations Research*, Vol. 18 No. 2, pp. 151-176.

Hon, L.C. (1998), "Demonstrating effectiveness in public relations: goals, objectives, and evaluation", *Journal of Public Relations Research*, Vol. 10 No. 2, pp. 103-135.

Jarzabkowski, P., Balogun, J. and Seidl, D. (2007), "Strategising: the challenge of a practice perspective", *Human Relations*, Vol. 60 No. 1, pp. 5-27.

Kim, M. and Cappella, J. (2019), "Reliable, valid and efficient evaluation of media messages: developing a message testing protocol", *Journal of Communication Management*, Vol. 23 No. 3, pp. 179-197.

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

Lindenmann, W. (1997), "Guidelines for measuring the effectiveness of PR programs and activities", available at: www.instituteforpr.org/wp-content/uploads/2002_MeasuringPrograms.pdf (accessed 8 December 2018).

Macnamara, J. (2014), "Emerging international standards for measurement and evaluation of public relations: a critical analysis", *Public Relations Inquiry*, Vol. 3 No. 1, pp. 7-29.

Macnamara, J. (2015), "Breaking the measurement and evaluation deadlock: a new approach and model", *Journal of Communication Management*, Vol. 19 No. 4, pp. 371-387.

Macnamara, J. (2018), "'Future directions for evaluation research', response delivered at the panel 'New Voices in PR Evaluation: innovative Approaches and New Research Avenues for a Field in Stasis'" *68th Annual International Communication Association Conference, Prague, 24–28 May*.

Macnamara, J. and Gregory, A. (2018), "Expanding evaluation to progress strategic communication: beyond message tracking and control to open listening", *International Journal of Strategic Communication*, Vol. 12 No. 4, pp. 469-486.

Macnamara, J. and Likely, F. (2017), "Revisiting the disciplinary home of evaluation: new perspectives to inform PR evaluation standards", *Research Journal of the Institute for Public Relations*, Vol. 2 No. 2, pp. 1-21.

Macnamara, J. and Zerfass, A. (2017), "Evaluation stasis continues in PR and corporate communication: Asia-Pacific insights into causes", *Communication Research and Practice*, Vol. 3 No. 4, pp. 319-334.

Marklein, T. and Paine, K. (2012), "The march to standards", paper presented at AMEC's 4th European Summit on Measurement, Dublin, 13–15 June, available at: http://amecorg.com/downloads/dublin2012/The-March-to-Social-Standards-Tim-Marklein-and-Katie-Paine.pdf (accessed 23 April 2019).

Michaelson, D. and Stacks, D.W. (2011), "Standardisation in public relations measurement and evaluation", *Public Relations Journal*, Vol. 5 No. 2, pp. 1-22.

Nothhaft, H. and Stensson, H. (2019), "Explaining the measurement and evaluation stasis: a thought experiment and a note on functional stupidity", *Journal of Communication Management*, Vol. 23 No. 3, pp. 213-227.

Nothhaft, H., Werder, K.P., Verčič, D. and Zerfass, A. (2018), "Strategic communication: reflections on an elusive concept", *International Journal of Strategic Communication*, Vol. 12 No. 4, pp. 352-366.

Place, K.R. (2015), "Exploring the role of ethics in public relations program evaluation", *Journal of Public Relations Research*, Vol. 27 No. 2, pp. 118-135.

Ragas, M.W. and Laskin, A.V. (2014), "Mixed-methods: measurement and evaluation among investor relations officers", *Corporate Communications: An International Journal*, Vol. 19 No. 2, pp. 166-181.

Romenti, S. and Murtarelli, G. (2018), "'Conflicting logics between communication departments and management' presentation delivered at the panel 'New Voices in PR Evaluation: innovative Approaches and New Research Avenues for a Field in Stasis'" *68th Annual International Communication Association Conference, Prague, 24–28 May*.

Romenti, S., Murtarelli, G., Miglietta, A. and Gregory, A. (2019), "Investigating the role of contextual factors in effectively executing communication evaluation and measurement: a scoping review", *Journal of Communication Management*, Vol. 23 No. 3, pp. 228-245.

Sandberg, J. and Tsoukas, H. (2011), "Grasping the logic of practice: theorising through practical rationality", *The Academy of Management Review*, Vol. 36 No. 2, pp. 338-360.

Schriner, M., Swenson, R. and Gilkerson, N. (2017), "Outputs or outcomes? Assessing public relations evaluation practices in award-winning PR campaigns", *Public Relations Journal*, Vol. 11 No. 1, pp. 1-15.

Shneider, A.M. (2009), "Four stages of a scientific discipline; four types of scientist", *Trends in Biochemical Sciences*, Vol. 34 No. 5, pp. 217-223.

Sommerfeldt, E.J. and Buhmann, A. (2019), "The status quo of evaluation in public diplomacy: insights from the US State Department", *Journal of Communication Management*, Vol. 23 No. 3, pp. 198-212.

Stacks, D.W. (2017), *Primer of Public Relations Research*, 3rd ed., Guilford Press, New York, NY.

Stufflebeam, D.L. and Coryn, C.L.S. (2014), *Evaluation Theory, Models, and Application. Research Methods for the Social Sciences*, 2nd ed., Jossey-Bass, Hoboken, NJ.

Swenson, R., Gilkerson, N., Likely, F., Anderson, F.W. and Ziviani, M. (2019), "Insights from industry leaders: a maturity model for strengthening communication measurement and evaluation", *International Journal of Strategic Communication*, Vol. 13 No. 1, pp. 1-21.

Thummes, K. (2009), *Ist Kommunikation messbar? Eine kommunikationswissenschaftliche Analyse der Quantifizierbarkeit von Kommunikation und aktueller Ansätze des Kommunikations-Controllings*, Helios, Berlin.

Toth, E.L. (2002), "Postmodernism for modernist public relations: the cash value and application of critical research in public relations", *Public Relations Review*, Vol. 28 No. 3, pp. 243-250.

Van Ruler, B. (2019), "Agile communication evaluation and measurement", *Journal of Communication Management*, Vol. 23 No. 3, pp. 265-280.

Volk, S.C. (2016), "A systematic review of 40 years of public relations evaluation and measurement research: looking into the past, the present, and future", *Public Relations Review*, Vol. 42 No. 5, pp. 962-977.

Watson, T. (1997), "Measuring the success rate: evaluating the PR process and PR programmes", in Kitchen, P.J. (Ed.), *Principles and Practice of Public Relations*, International Thomson Business Press, London, pp. 293-294.

Watson, T. (2012), "The evolution of public relations measurement and evaluation", *Public Relations Review*, Vol. 38 No. 3, pp. 390-398.

Wehmeier, S. (2006), "Dancers in the dark: the myth of rationality in public relations", *Public Relations Review*, Vol. 32, pp. 213-220.

Wehmeier, S. and Nothhaft, H. (2013), "Die Erfindung der 'PR-Wissenschaft': Bemerkungen zu Theorie und Praxis und Wege aus der Delegitimierungsfalle", in Hoffjann, O. and Huck-Sandhu, S. (Eds), *UnVergessene Diskurse. 20 Jahre PR- und Organisationskommunikationsforschung*, VS Verlag für Sozialwissenschaften, Wiesbaden, pp. 103-134.

Werder, K.P., Nothhaft, H., Verčič, D. and Zerfass, A. (2018), "Strategic communication as an emerging interdisciplinary paradigm", *International Journal of Strategic Communication*, Vol. 12 No. 4, pp. 333-351.

Whittington, R. (2007), "Strategy practice and strategy process: family differences and the sociological eye", *Organisational Studies*, Vol. 28 No. 10, pp. 1575-1586.

Zerfass, A., Verčič, D. and Volk, S.C. (2017), "Communication evaluation and measurement: skills, practices and utilisation in European organisations", *Corporate Communications: An International Journal*, Vol. 22 No. 1, pp. 2-18.

Further reading

Ihlen, Ø., Gregory, A., Luma-aho, V. and Buhmann, A. (2019), "Truth, evaluation and education", *Public Relations Review*.

Ingenhoff, D. and Buhmann, A. (2016), "Advancing PR measurement and evaluation: demonstrating the properties and assessment of variance-based structural equation models using an example study on corporate reputation", *Public Relations Review*, Vol. 42 No. 3, pp. 418-431.

Van Ruler, B. (2015), "Agile public relations planning: the reflective communication scrum", *Public Relations Review*, Vol. 41 No. 2, pp. 187-194.

Watson, T. and Noble, P. (2014), *Evaluating Public Relations: A Guide to Planning, Research and Measurement*, 3rd ed., Kogan Page Publishers, London.