Theoretical Sampling – ALIAS: a Case selection framework for research on social media engagement

M. Wolf, J. Sims, H. Yang

Introduction

Social media is a new and emergent phenomenon. Its supports interaction and participation in the generation of content, and facilitates the emergence of a new ideology of open access and collaboration (DesAutels, 2011). The content which was traditionally created and controlled by corporate content providers (e.g. product information, encyclopaedia, phone register) is now being co-created by the consumers (DesAutels, 2011; Kietzmann, Hermkens, McCarthy, & Silvestre, 2011).

Social Media allow different groups of people to engage in conversation and so to exchange ideas and generate new content, services and products. Many firms try to engage with their stakeholders via social media, some are successful, others less so. How can better and worse performing firms be identified and isolated for further investigation?

Open public access to many of the social media sites such as Facebook, twitter, LinkedIn, and others, allows an observer to sense the sentiment of exchange, to gauge the level of engagement, and to see who is taking part in a conversation. The what is happening can be observed, however, the how and why – how can some firms create engagement and why do some firms fail to do so – remains unseen. Acknowledging that all participants (that is all firms seeking engagement) have equal access to the media, the successful engagement is rooted in “unique and idiosyncratic resources and capabilities” (Rouse & Daellenbach, 1999, p. 488). This paper therefore, adopts a Resource Based View (RBV) of the firm (Barney, 1991) and focuses on qualitative case study research in organisations with the aim to uncover firm specific capabilities, policies and processes of successful organisational engagement.

This paper starts by framing the notion of dialogue and putting it into the context of firm-employee engagement on social media. The definition of theoretical sampling is then introduced. The core of the paper is the introduction of a Case-Selection Methodology followed by an example of how this methodology was applied in a PhD research project, and conclusions.

Dialogue and DART

This paper focuses on the conversations between an organisation and its stakeholders on social media. Such an engagement can contribute to (co-) creation of value for all parties (Kaplan & Haenlein, 2010). (Prahalad & Ramaswamy, 2004) summarise the features of a co-creation system in their DART-framework (Figure 1), attributing co-creation value to *Dialogic* communication, *Transparent Access* to information by all parties, and *Risk-benefits* balance (added value for all participants).

![Figure 1 - DART-Model from Prahalad and Ramswamy's (2004)](image)

A *dialogue* is conversation in which a power balance between all participating parties is maintained (Kent & Taylor, 1998). The idea of the balance of power goes back to Habermas’ ideal speech situation which requires all parties to be true to themselves, have the same opportunity to participate and equal power to influence others, and also allows every statement to be questioned and debated (Leeper, 1996). The power according to (Lukes, 1974) has three faces: the power to speak, the power to define what can be said and the power to prevent others from speaking.

Linking these power attributes back to DART (Figure 2) – *dialogue* requires an equilibrium of powers and contributes to it; equality of *access* contributes to distribution of power and is influenced by power shifts at the same time; *risk-benefits* balance influences participants’ decision to speak or not to speak and to challenge and debate decisions; *transparent* information contributes to power distribution and informs the participants’ contribution.

![Figure 2 - Linking DART framework to Power](image)
Who speaks and actively participates in conversation
What is being said, and
Who is excluded

The better performing organisations would have a higher level of dialogue and access (many and diverse active participants), high transparency (what is being said) and few exclusions. The justification and method for the identification of “best performers” is guided by the idea of theoretical sampling (Creswell, 2013; Eisenhardt, 1989; Rouse & Daellenbach, 1999) for theory building and is discussed in the following sections.

**Theoretical sampling**

The differences in the assumptions about the world and knowledge (ontology) and valid ways of obtaining this knowledge (epistemology) find reflection in approaches to theory building. On one side of the spectrum middle-range-theory (MRT), endorsed by (Merton, 1957), focuses on inferring relationships between pre-conceived variables and creating theory by putting them “to the test of observation by seeing whether these inferences turn out to be empirically so” (Merton, 1957). On the other side, Grounded Theory, famously introduced by Glaser and Strauss (1967), focuses on the theory-building with major emphasis placed on qualitative data and flexible or emergent research design (Layder, 1993). Theory building process begins as closely to the “ideal of no theory under consideration” as possible (Eisenhardt, 1989, p. 536), the researcher enters an iterative process of collecting and analysing data during which the constructs emerge. The theoretical constructs, unlike MRT, and the relationships between them become apparent during the data-analysis and are not preconceived (Eisenhardt, 1989). The data collection is guided by the principle of “theoretical sampling” – the most revealing or outstanding cases are selected for closer investigation (Creswell, 2013). This means, that instead of selecting a possibly random and statistically representative sample, a few “relevant” cases are selected (Eisenhardt, 1989; Rouse & Daellenbach, 1999; Yin, 2009). The general approach is to select (1) theoretically relevant cases and (2) as many as required to reach saturation. While (Yin, 2009) suggest to limit the number of cases to 6-10, arguments for a smaller or indeed greater number of cases based on theoretical saturation, predicted replicability or contrast of cases can be found (Eisenhardt, 1989; Rouse & Daellenbach, 1999). The theoretical saturation is reached when every new case is not revealing any more new data, insights and constructs. Knowing the number of cases prior to data collection and data analysis is therefore difficult and the initially planned number might need to be adjusted.
Theoretical relevancy is driven by two factors. First, the selection of relevant cases allows a certain level of control for environmental factors, for example selecting firms from the same country, industry and of similar size (Eisenhardt, 1989). Second, the selection of cases is driven by the desire to select cases which are likely to yield insights. The focus is not on a “typical” case (as it would be with a representative sampling for statistical analysis), but rather on a “telling” case which is likely to “make previously obscure theoretical relationships apparent” (McKeown, 1999, p. 174), or by the approach which (Levitas & Chi, 2002) critiqued as “sampling on the dependent variable” (p. 961) – selection of cases based on the observable “outcome”, for example most successful companies, most popular blogs, best paid actors etc.. Each case in multi-case study design represents a single case study in itself, so that the consideration for selection of cases for the single-case design apply here too. However, in addition to single-case considerations, the cases are chosen in conjunction with each other. For example extreme cases of polar types, or similar cases with controlled environmental variation (Eisenhardt, 1989; Glaser & Strauss, 1967; Rouse & Daellenbach, 1999). One of the frameworks to select cases has been introduced by (Rouse & Daellenbach, 1999). The four step selection process starts with identification of the industry and collection of relevant performance data. The firms are then grouped based on a multi-facettied list of strategic attributes into groups or clusters. In the third step, the key performance indicators of group members are compared and, finally, high and low performers are selected for closer inspection (Rouse & Daellenbach, 1999). In order to be able to identify clusters, performance indicators and select high- and low performers some a priori constructs are necessary (Eisenhardt, 1989).

**ALIAS – a Case Selection Methodology**

While the “what” on social media is publicly available and observable (who are participants, who create the posts, who comment on them and what is the content) a large sample study is unlikely to reveal why and how some firms manage to create a dialogue while others don’t. The proposed framework aims at aiding the selection of *theoretically relevant* cases, i.e. the identification of “high and low performers” (Eisenhardt, 1989; Rouse & Daellenbach, 1999) in the context of dialogic social media engagement. To clarify the framework, the example of a hypothetical research setting, investigating co-operative engagement between software developers and corporate customers based on their participation in a User-Group on a fictive social media site is used.
The case selection framework introduced in this paper: ALIAS – is a five step process for identification and purposeful selection of case-study cases (Figure 3).

![Figure 3 - ALIAS - steps of the case selection process](image)

**Actor Identification**

The selection process starts with the identification of (possible) actors, e.g. developers/users. The a-priori definition of at least some of the participants is important to address the question “who is excluded”. Notably, the (possible) actors can and probably will change during data collection phase (e.g. business analysts might emerge as a distinctive participant group); the preliminary list is used as a guidance during the case selection process only.

**Limiting the population**

In the second step, the population of potential cases is defined, e.g. UK B2B software firms. This aids a) the limitation of the number of cases, and b) controls for environmental variations (Eisenhardt, 1989; Glaser & Strauss, 1967; Rouse & Daellenbach, 1999). It further sharpens the focus of the study and increases potential for transferability of findings (Denzin & Lincoln, 2008).

**Identification of Observables**

The third step identifies desired “observables” – the performance indicators on social media, e.g. comments and replies in the user-group. This step is mirroring (Rouse & Daellenbach, 1999) identification of “performance data” (p. 489). It might be necessary to include the definition of measurements for each of the indicators. While some indicators are binary (e.g. “posting on face book enabled / disabled); others are quantitative (e.g. average number of re-tweets, number of comments per post, number of video views); and yet some are qualitative (e.g. sentiment of product reviews; video content). The measurement scales are then applied in the Assessment-step.

**Assessment**

During this step, participant’s engagement and activity levels are assessed based on the indicators identified previously. Various indicators are then compared across all cases and a value is assigned to the actor-performance to indicate their relative performance. The resultant matrix allows arrangement of cases based on each actor-group’s performance.
Figure 4 - Performance Matrix – demonstrates an example of comparative engagement of participants from different firms and highlights the theoretically promising cases.

In the example used here, the evaluation of performance indicators is fairly simple – wither a simple “comment/do not comment” or a count of the number of comments would yield enough data to allow such an arrangement. Real life examples are much more complex than simple “comment/do not comment” on one distinct platform; in the next section a more complex assessment of performance indicators is discussed using an ongoing PhD research project as an example.

The qualitative data, as will be demonstrated in the next section, can either be quantified, or be used by the researcher to adjust the positions. The framework, true to subjectivist spirit, is intended as a guideline and does not claim universal prediction powers.

Selection of cases

Once the arrangement of firms is completed, the best/worst performers can be visually (and quantitatively) identified. In the example Figure 4 “Theta Inc” and “BB-Soft” appear to create much higher levels of engagement than others; “XetaDev” appears to have actively participating developers, but disengaged customers, whereas “C-Industries” customers appear to be much more active than its developers. “Omega Corp” presents another interesting case, because no engagement could be observed – one of the questions to ask: is there really no engagement or was it an error in observation? The selection of “promising” cases is still the task of the researcher, with more confidence and guidance from the assessed performance data.

Application of the ALIAS framework

To identify the firms who successfully (or less so) engage with their employees on social media, the ALIAS framework has been applied during the case selection for a PhD research project which aims to understand how HRM as a strategic discipline addresses the challenges posed by social media by juxtaposing best and worst performers (Rouse & Daellenbach, 2015-20150118 v2 js).
The study focuses on internal, idiosyncratic and firm specific processes, practices and policies.

**Social Media in HRM – theoretical sampling**

While Human Resource Management (HRM) favours a top-down, strategic approach, social media is an emergent, bottom-up phenomenon. Integration social media into HRM therefore poses potential challenges. A PhD research project, aimed at understanding how firms deal with this challenge, investigates firm specific strategies, policies and practices in relationship to social media use. The application of the ALIAS framework in the process of theoretical sampling for this study is presented thus.

**Social Media as User Generated System**

Social Networking Sites (SNS) are web-based applications built upon the technological foundation of Web2.0; many of them allow users to not only generate content, but also to explicitly express their identity and their relationships with each other (Boyd & Ellison, 2008). Social media is a user generated information system which integrates SNS, Web2.0 and other technologies and provides unique value to the user (DesAutels, 2011). The ideological foundation of Web 2.0 is rooted in the open source ideology, whereby users have free access to information and tools and can create and expand the available resource base in collaboration with other participants (Boyd & Ellison, 2008; DesAutels, 2011; Hauptmann & Steger, 2013; Kietzmann, et al., 2011). The ideas of open access, open source and collaboration make social media inherently “bottom-up” and democratic.

**Human Resource Management as strategic discipline**

The main research streams in HRM are focused on establishing linkages between HRM and organisational performance (Guest, 2011 #37). (Guest, 2011) identifies three different directions from which these linkages were investigated: one focusing on HR practices, another applying the resource based view to HRM, and lastly a focus on implementation of a set of HR practices. The common denominator of these approaches is the search for the source of competitive advantage (Barney, 1991; Guest, 2011; Miles & Muuka, 2011; Wright, Dunford, & Snell, 2001). (Guest, 2002) identifies three key models which link HRM to improved organisational performance: High Performance Work System; High Commitment; and the Strategic Fit model. A rather recent addition to this list is “process view” introduced by (Bowen & Ostroff, 2004). All of these theories agree on the strategic position of HRM and focus on strategic top-down vertical and horizontal alignments of HR practices, policies and strategies (Boxall & Macky, 2009; Collings & Mellahi, 2009; Saks, 2006).
This paper adopts the lens of RBV and asks, what are the firm specific factors what allow organisations to successfully integrate social media in their HR activities.

**Actor identification**

Two pilot studies were conducted early 2013 in large UK organisations. The findings revealed that firms seem to engage with their employees on social media prior to the employment (candidates and applicants), during the employment (employees) and after the employment (alumni). Adopting the terminology used by the informants in the pilot studies, for the selection of cases all employees (former, current and future) are referred to as “employees”. So that the two broad groups of actors identified are Employees (acting as individuals) and Firm (acting as an official entity). The distinction within the “employee” – actor group is, however, essential for identification of observable outcomes. The creation of company pages on SNS, posing of comments and replies on in the name of the firm, etc. is, arguably, still done by individual employees, however, these employees are acting distinctly on the behalf of the “organisation” and not as individuals.

**Limiting the population**

The population was limited in three steps. Each step addressed one specific issue and helped sharpening the focus of the research.

*Step 1 – Only those who do*

The first issue addressed was that of “non-engagement”: if a firm and its employees are not seen to engage on social media, is it because they actually do not, because the researcher is not looking in the right place? How does one observe something which is not there? To address this problem, the initial population of organisations to be reviewed was limited to 408 organisations who participated in the London Organising Committee for Olympic Games (LOCOG) social network during the outplacement of LOCOG employees after the London 2012 Olympics was over. All four hundred organisations did engage on SNS at least once during the London 2012 Olympics. Arguably, if one of these firms was not present on any public SNS, did not link to any SNS from their homepage and careers page, it could be assumed that this organisation is consciously not actively engaged on public SNS.

*Step 2 – Only those who can*

Many of the organisations were small and relied on external support to manage their engagement. LOCOG’s network allowed employers to either target individual employees directly, or to set up groups and engage in more general discussions with a broader population of employees. Less than one hundred firms engaged with the employees in this way. Some others used recruitment or recruitment process outsourcing agencies to taken on
this role; these companies were excluded – the research focuses on direct communication between the firm and its employees, without facilitation of third parties. The remaining list contained just over fifty firms most of whom had more than 25,000 employees, although some of the engaged businesses employed as little as 5,000 people.

**Step 3 – Only those who are accessible**

Finally, the list was reduced to thirty nine UK based companies of which 32 had more than 25,000. It seems an unachievable target to conduct a case study in a firm based in Rio de Janeiro or in Moscow – the limitations of time, money and language barriers have to be accepted. In addition, limiting the population to the UK allows to control for environmental factors (Eisenhardt, 1989; Rouse & Daellenbach, 1999) such as political climate, legislation, workforce education levels, unions etc.

**Identification of Observables**

The observables for each of the actor-groups differ somewhat; in their definition the DART framework was used as a guideline, albeit not all elements of the DART framework were applicable to each group. The pilot studies revealed that four SNS were predominantly used by HR employees – LinkedIn, twitter, YouTube and Facebook. Other SNS, notably Google+, Pinterest and Glassdoor were uncovered during the assessment step, however the activity on those sites was not evaluated during the assessment step.

**Observables for Firms**

*Dialogue* was considered to take place when the firm was actively replying to posts or comments on LinkedIn and Facebook, replying or re-tweeting on twitter, or commenting on YouTube. *Access* was considered to be granted when posting was enabled on Facebook, reviews enabled on Facebook, following was possible on twitter, comments enabled on YouTube channel and videos, careers and Alumni groups were open on LinkedIn. *Transparency* of information was considered to be present when posts, tweets and videos carried more than corporate message and job-postings. In addition, the ease of access across the platforms (i.e. extant cross-links between the company home page and SNS) contributed to transparency.

**Observables for Employees, Candidates and Alumni**

*Dialogue* was considered to exist when employees posted updates, videos or comments on any SNS. *Risk/Benefit* value was considered higher if the comments were critical, or posed questions, suggesting that the information would benefit the employee (for example candidates asking recruiters, or an ex-employee critiquing the firm). *Transparency* was judged high, when the names or relationship to the firm were exposed. In addition, exchanges
within employee groups, e.g. candidates / alumni or current employees/alumni were considered to contribute to transparency

**Assessment**

The complexity of the actor-group composition and the number of observables posed a challenge. The firm actors were assessed in two dimensions: based on their (1) encouragement/discouragement of engagement, and on their (2) active participation/non-participation; the employee actors were assessed in two different dimensions: (1) their engagement/non-engagement with the firm actors, and equally (2) engagement/non-engagement within the actor-group. Not only was the amount of information to be collected very high, the data formats were a mix of qualitative and quantitative data. The question “can employees post comments on YouTube” is answered with a “yes/no” or 1/0 by just looking at the site, whereas “do videos on YouTube carry an authentic and transparent message?” is difficult to answer even after watching a number of those. The above criteria (seventeen in total) were grouped corresponding to the assessment dimensions: (1) firm encourages participation (enabled comments, allows postings), (2) firm engages (actively posts, replies), (3) employee engages (posts, replies), (4) employee engages outside firm’s SNS (alumni groups, conversations outside firm posts). Each firm could score 10 points and Employees 7 points (plus any additional points granted). The firms where then sorted on their scores into four quadrants (Figure 5).

Firms with higher scores were positioned closer to the corresponding corners, with firms close to average located near the centre. In addition to the pre-defined observable performance indicators, reflective and subjective comments were written down next to each set of scores. These were used to support decision making during the selection step.

**Selection of cases**
Based on the results of the assessment steps, 10 “most interesting” firms have been selected. These included five firms from the square 1 (above average firm and employee engagement), three from the square 4 (above average firm, below average employee engagement) and two from the square 2 (below average firm, above average employee engagement). Square three has been ignored: the study aims at understanding at why some firms are successful in building engagement (square 1), whereas other try to build engagement and are less successful (square 4) or do not try to create engagement when they could (square 2).

All HR Directors, Heads or Recruitment and LinkedIn-group owners were contacted with details of study and a permission to conduct a study in their organisation requested. At present three organisations (all from square 1, not surprisingly) have replied and displayed interest in further study.

**Conclusion**

Entangling firm specific idiosyncratic relationships enables researchers to develop insight into *why* and *how* sustained competitive advantages can be achieved (Barney, 1991; Rouse & Daellenbach, 1999). (Rouse & Daellenbach, 1999) call for research in organisations and uncovering the *how* successful firms made themselves different from others. The challenge, however, is the identification of “successful” firms (and correspondingly the not so successful) for juxtaposition and comparative study (Eisenhardt, 1989). This paper presented a five step ALIAS framework that guides identification of theoretically relevant cases (Yin, 2009) and demonstrated the application of this framework in a selection of cases for study of social media use in HRM in large UK firms.

The challenges of case selection and identification and measurements of performance criteria on social media have been highlighted and discussed.

**Limitations**

Further theorising on how performance criteria can be reliably measured will contribute to sharpening of constructs and improved reliability of case classification. The aim of this selection process at this stage is to guide the researcher in case selection and not to provide a definitive fixed set of cases to be studied.

Whilst the paper suggest that the extreme, corner cases are the theoretically relevant cases, it needs to be acknowledged, that comparison with “normal” cases, those closer to average performance could also benefit the research and formulation of theory.

The reduction of a multi-dimensional space (in the context of the example study a four dimensional space of (1) firm’s encouragement of dialogue, (2) firm’s participation in
dialogue, (3) employee’s participation in dialogue, and (4) employee’s construction of own
dialogue) has been reduced to two dimensions. The complex and diverse actor communities
have been reduced to just two “generic” types (e.g. HRM, management, Public Relations,
Marketing etc. are grouped as “organisation”-actor; experienced candidates, graduates,
employees at all levels, alumni etc. are group as “employees”). This reductionism allows to
keep the assessment model simple and the taxonomy accessible, at the same time a balance
between complexity and detail might lay in a more sophisticated assessment model (such as
the 17-criteria model used in the study).

**Contribution**

This paper has academic and practitioner implications. First, the academic community will
find the selection framework helpful in guiding and justifying selection of cases for in-depth
studies. Unlike (Eisenhardt, 1989) suggestion to use qualitative studies to build quantifiably
testable theories, this approach uses positivist quantifiable data to identify opportunities for
qualitative research. Second, the step-by-step approach to case selection helps a gradual
reduction of cases and addresses the concern of being “drown by the data” (Creswell, 2013;
Eisenhardt, 1989; Yin, 2009) – in this case the number of potential research sites. Third,
academic researchers will find that the rigour and transparency of the selection procedure
improves reliability of their selection process and contributes to transferability of later
findings (Denzin & Lincoln, 2008). Fourth, the framework allows academic and practitioner’s alike to identify actor communities of social media use. Such an identification
contributes to clarity of DART-features: the expected Dialogue, Access routes and media,
Risk/Values of social media use, and identification of data and information to be made
transparent. Fifth, identification of measurable outcomes or effects (performance data)
supports academics and practitioners in establishing success criteria and only then allows a
like-for-like comparison of organisations with the aim of identification of higher- and lower
performers. Such an identification would assist academics in the selection of cases, and
practitioners in identification of areas for improvement.

**References**

Management, 17*(1), 99-121.

Role of the “Strength” of the HRM System. *Academy of Management Review, 29*(2),
203-221.


