This paper is from the BAM2019 Conference Proceedings

About BAM

The British Academy of Management (BAM) is the leading authority on the academic field of management in the UK, supporting and representing the community of scholars and engaging with international peers.

http://www.bam.ac.uk/
ACCESSING THE INACCESSIBLE: E-SAMPLING VIA FACEBOOK

Iuliana Chitac, PhD Candidate, University of Westminster
w1654057@my.westminster.ac.uk

Dr. Deborah Knowles, University of Westminster
D.S.Knowles@westminster.ac.uk
Abstract

**Background:** This paper presents accessing and sampling strategies, including e-sampling via Facebook, and their time efficiencies, as used by an interpretative PhD study focused on the “hard-to-access” community, of London-based Romanian entrepreneurs.

Although, statistically, Facebook is the most popular social media platform, its research potential is under-utilised by researchers. Despite being considered a barrier-breaking research tool, particularly when recruiting hard-to-reach populations, its research potential has been mostly explored for surveys and adverts, rather than as a recruitment platform for qualitative interviewing. Additionally, a comparative assessment of time and informants’ participation rate efficiencies associated with different sampling techniques has been largely overlooked.

**Objectives:** To address these gaps, the following questions have been formulated, given the limited PhD resources and the access barriers to “hard-to-access” populations: does combining traditional sampling with e-sampling via Facebook increase research informants’ selection? Is e-sampling via Facebook on average more time efficient compared to other traditional sampling techniques?

**Methodology/Approach:** To answer these questions, an e-sampling process via Facebook was designed to recruit London-based Romanian entrepreneurs participating in face-to-face interviews. Additionally the estimated average time it takes to successfully sample an interviewee for each sampling technique used in this qualitative PhD study was calculated.

**Findings:** E-sampling via Facebook enjoyed a similar participation rate as the traditional snowballing and it was the most time efficient (2:32hrs per interviewee), playing a critical role in research selection, particularly when traditional selection techniques dried up.

This paper contributes to knowledge by offering context-bounded insights on how to efficiently access and sample hard-to–reach informants through e-sampling via Facebook. It provides new evidence of average time necessary in selecting research participants for qualitative interviewing using different sampling techniques.

**Keywords:** Hard-to-access population, entrepreneurs, e-sampling, Facebook, qualitative research, PhD.
1. Introduction

When conducting qualitative research, which requires gaining access and selecting “hard-to-reach” communities, such as immigrants, ethnic and sexual minorities and women (Cundiff, 2012; Ellard-Gray et al., 2015), researchers should adopt a reflective and a flexible, iterative approach (Hoppitt et al., 2012). This could improve the selection process and the quality of the research (Cunliffe and Alcadipani, 2016; Mikecz, 2012). This approach could enable the identification of effective and efficient research solutions to deliver a rigorous study (Thornton et al., 2016), in a challenging research context, where accessing and sampling in these communities means overcoming low visibility and location (Shaghaghi et al., 2011; Sydor, 2013), social stigma (Hooker, 2010) and cultural and language “otherness” (Marpsat and Razafindratsima, 2010).

Although the time, budget and selection efficiencies enabled by conducting research using social media are undeniable (Bik et al., 2015), and Facebook’s increasing influence, reach and power are statistically reinforced, its potential as a research tool remains underexplored (Kosinski et al., 2015) compared to Twitter and blogging (Bik et al., 2015; Bik and Goldstein, 2013). It continues to be explored only by a handful of scholars, mostly from the medical (Arcia, 2014; O’Conner et al., 2014) and social science fields (Kosinski et al., 2015). Despite being considered a barrier-breaking research tool, particularly when selecting hard-to-reach populations (Kosinski et al., 2015), its research potential has been mostly explored for surveys (Baltar and Brunet, 2012; Valdez, 2014) and adverts (Carter-Harris et al., 2016; Potzchke and Braun, 2017), rather than as a recruitment platform for qualitative interviewing. Additionally, a comparative assessment of time and informants participation rate efficiencies associated with different sampling techniques has been largely overlooked (Topolovec-Vranic and Natarajan, 2016).

Considering the particularities of hard-to-reach populations, social media might provide new, cost-effective research solutions (Rife et al. 2016). This approach is reinforced in the context of migration; Facebook plays an important role in migrants’ everyday lives, as means to communicate with their home country (Dekker and Engbersen, 2014) and with their nationals in the host country (Oiarzabal, 2012).

Given that on Facebook the limit between private and public information is not straightforward, using this data could raise ethical issues. These issues are more likely in quantitative research, when data is collected without the active participation of the informants and thus research consent is not an established practice (Golder et al., 2017). Acknowledging that boundaries between public and private do not stem from accessibility, but from the use of this data (Kosinski et al., 2015), in the context of this study no data apart from the consented interviews was used. Additionally, the researcher asked permission from the research participants themselves to access their Facebook network for research purposes and she always prioritized research transparency by respecting informants’ privacy and consent decision (Solberg, 2010). The researcher adopted “data protection by design and default” throughout the data collection process, by restricting the amount of data collected to research objectives, pseudonymising the recorded data, passwording the data and by using university’s cloud storage (Information Commissioner Office, 2019), as reinforced by GDPR (Article 25, 2018) through Ethics Committee of the University of Westminster.
This article’s contribution to knowledge is three-fold.

Firstly, it offers context, case study bounded accessing strategies and sampling techniques from an interpretative qualitative PhD study.

Secondly, no previous qualitative study has been found which provides evidence of the average time spent recruiting hard-to-access participants per sampling technique used, including esnowballing, derived rapport, snowballing and time-space. The PhD researcher designed a step-by-step virtual snowballing sampling process via Facebook, which enabled accessing and motivating 13 research interviewees within only 3 weeks between 19th November and 9th December 2018.

Thirdly, this article prioritises transparency and reporting of qualitative research experiences in publishing, as a venue to increase research quality, rigour and collaboration. These practical recommendations from researchers could help align the empirical reality with the qualitative research benchmarks formulated in Consolidated Criteria for Reporting Qualitative Research (COREQ) (Tong et al., 2007) and the Standards for Reporting Qualitative Research (SRQR) (O’Brien et al., 2014). This enables knowledge creation and dissemination in an efficient manner across a broad spectrum of researchers (Ellard-Gray et al., 2015; Rockliffe et al., 2018).

2. PhD researcher’s outsider/insider positionality

These reflective experiences stem from an interpretative, intersectional qualitative PhD study focused on Romanian immigrant entrepreneurs (RIEs) in London, conducted at the University of Westminster, UK.

The PhD researcher shares cultural and migration status background with the researched community, which helped establish her positionality as community “insider”. However, she also self-identifies as an “outsider” to this community since she is not an entrepreneur.

The next three sections review the literature on accessing and sampling hard-to-access communities, with a focus on immigrant entrepreneurs, whilst integrating the experiences of the PhD researcher on this study. Following this, the results of the accessing process are presented as barriers and strategies, culminating with detailing the e-sampling process via Facebook and with a comparative assessment of the estimated average time spent in recruiting per research participant for each sampling technique used.

3. Love you, love you not: barriers and strategies to gaining access to hard-to-reach communities

Human beings and their experiences are at the heart of qualitative research (Jacob and Ferguson, 2012). Their time and generosity in entrusting and sharing their lives should not be taken for granted. Reciprocal respect, ethical approach and quality should be carefully reinforced at all stages of the research. The journey of knowledge creation should be win-win for all involved.
When researching “hard-to-access” communities, including immigrants, the research process is often hindered, delayed or even deemed unachievable by practical and ethical challenges in accessing and sampling them (Ellard-Gray et al., 2015; Sanjari et al., 2014). Access barriers, including expectations of socio-cultural power dynamics, research participation risks, lack of trust in research impact and experience of failed ethics (Bonevski et al., 2014) made this research very time consuming, pushing the limits of what is achievable at times.

Most of the research on accessing hard-to-reach communities presents access in a linear way, as a sampling issue, reflected by the number of research participants or as a researcher agency issue (Freeman, 2000, cited in Cunliffe and Alcadipani, 2016). The PhD researcher avoids this arguably faulty and unidirectional approach, by emphasising the co-creative, collaborative nature of the researcher-research participant relationship that defines this phenomenological study, from sampling to findings. In qualitative research, gaining access should be an iterative, reflective, trust building and dynamic process (Driver, 2016; McDonald, 2015). This sets up the foundation for meaningful research engagement, contextual nuances and participants’ particularities to emerge in a natural, trusting way (Cunliffe and Alcadipani, 2016). The quality of this access drives the quality of knowledge co-creation enabled by how the researcher and the research participant relate to each other.

Similar to other qualitative studies focused on hard-to-reach populations (Baltar and Brunet, 2012; Bryman, 2012), gaining and maintaining access to Romanian immigrant entrepreneur community in London has been a challenging, dynamic and trust building process, spreading over nine months (April 2018 to January 2019). Initially, the PhD researcher engaged in an exploratory access, whereby the community of informants was observed (April-May, 2018). Secondly, during the trust-building access stage (June -January 2019), the researcher socially interacted with community members during community events. Consequently, online friendships and community membership materialised (Carmel, 2011), which enabled visibility of community events and identification of gatekeepers. Finally, during the breakthrough access stage (September- January 2019), the researcher became a recognised cultural insider and a valued and trusted community associate (Blix and Wettergren, 2015; Rantatalo et al., 2018).

During this interpretative qualitative PhD study, the researcher overcame varied access barriers, amongst which the most impactful were captured. Aware of the informants’ cultural importance of valuing relational approach to collaboration, whereby mutual respect, trust and accountability are prioritized about research and professional practices (Cunliffe and Eriksen, 2011), the PhD researcher used varied strategies to overcome the access barriers as detailed below.

**Researcher’s partial outsider positionality**

This influenced the members’ initial lack of openness and trust in this study. The perceived “otherness” of the researcher built suspicions around the declared research interests and benefits for the researched community (Zickar and Carter, 2010). Although the PhD researcher self-identifies as a Romanian immigrant, so a cultural and language insider, she was an outsider to the community she was researching.

To overcome this access barrier and thus to enable a relational access, the PhD researcher participated in community events and networks to increase familiarity and opportunities of detailing the research benefits for the community (Marshall and Rossman, 2011). Additionally,
by nurturing a relational collaboration with key community members and thanks to her shared cultural and language background, the PhD researcher became an associate member of the researched community (Barley, 2011). These allowed for direct communication with the potential research participants in their native language.

Within this context, the PhD researcher experienced that building a long-term, trustful relationship with the research community was vital, whilst establishing herself as a trusted member of this community.

Additionally, the PhD researcher identified the community “influencers” by accessing media trusted sources and she attended local events where these influencers were promoted. These events were opportunities to introduce the research and formulate an informal invitation to the research.

The researcher-participant and researcher-gatekeeper power relation barriers

These relationships are critical expressions of the social and professional status shaped by the socio-cultural context (Riese, 2018). Power is the way through which we influence our own and others’ behaviours during social interactions (Foucault, 1982). As creator of the research, the researcher was perceived as having superior positionality to research participants and gatekeepers, controlling the interview and the participants’ selection. This initial power relation might just turn off research participants for whom relational access, based on trust, is a prerequisite for collaboration (Cheek, 2011).

To overcome the power dynamic barrier, the PhD researcher reflectively adjusted the power relation that would be equally conducive of ethical and quality research on a one-to-one basis. She always tried to strike the right balance between the recognition given to the researcher’s professional skills and the quality of the research, whilst allowing for their professional status and the impact of their participation to be overtly acknowledged. Additionally, the researcher tried to prioritise the informants’ site, time and travel expenses or language preferences (Wendler et al., 2006, cited in Riese, 2018).

Lack of trust in the research process and its value-adding benefits for the researched community.

This mistrust was based on participants’ lack of research culture in their home country where research might lack transparency of practical benefits (Jacklin and Kinoshameg, 2008), or where ethics might be only loosely applied (Bonevski et al., 2014).

Research participants’ social and professional risks, including loss of professional reputation or social status if during the interview, for example, they reveal sensitive business practices (Bonevski et al., 2014).

To overcome these barriers, the researcher reinforced her association with the University and provided them with University’s Ethics Committee approved research documentation, which emphasise the GDPR rules. The documents of voluntary consent and anonymity were explained and signed before the interview. The transparency of these practices was reinforced by inviting participants to University of Westminster.
The research informants’ avoidance of pro-bono research participation, stemmed from research informants’ perceived professional differences between them and the researcher, being only conducive of an instrumental relation. A researcher is often perceived as being a theoretician with no practical ability for a fruitful, practical collaboration with real-life professionals like the entrepreneurs for example. Within this context, the informants may ask themselves: “if one hand cannot wash the other, what is the purpose of engaging in this research?” (a well-known Romanian expression) and they avoid collaboration. To overcome the perception of pro-bono research participation, the researcher “befriended” the research participants on Facebook to increase visibility of her multi-industry, global professional experience as a financial analyst at Morgan Stanley and a liquidation manager in the USA, captured at her profile. This enabled the identification of common interests between researcher and informant and further encouraged a relational approach, whilst creating the opportunity for win-win collaboration.

Research participant’s refusal to refer another informant from the researched community either to avoid asking favours they need to return later on, or because they perceive their research participation as an incremental competitive advantage and a recognition of their social status and thus they prefer to keep it this way.

To overcome this access barrier, the researcher emphasised the importance of sample size and quality under anonymity and confidentiality conditions (Vuban and Eta, 2018). Additionally, the researcher reflectively combined different sampling techniques to overcome the gatekeeper dependency and reduce the selection bias specific to sampling techniques.

Additionally, the researcher also identified the community’s most “connected” people during community events and negotiated with them to engage as “gatekeepers”, and thus to facilitate the participation of others in this research (Pritchard and Symon, 2014). Furthermore, proactively managing the possibility of research disengagement from these gatekeepers, she asked their permission to gain access to their online network, which included members of the researched community. This was used as an excellent opportunity for online sampling, which proved critical, particularly when traditional sampling stopped working and the gatekeepers were inefficient.

Gatekeepers have been crucial for this study as they were the ones who controlled the access to these research communities as influential people, community role models (Clark, 2011; Johl and Renganathan, 2010). The process of negotiating access to this community has been tedious and time-consuming. Gaining access to this community meant gaining access to multiple clusters, which were driven by strongly gendered enacted power relations, under the suspicion of “being played” and “being watched”. These social behaviours are deeply rooted in their patriarchal and autocratic upbringing and to overcome them it required a laborious one-to-one approach, whereby, each of the research informants with all their requirements became the central focus of the researcher.

It is important to remember that cultural background and past experiences shape our social behaviours as much as the context and our perceived similarity to or otherness from the people with whom we are interacting. Through reciprocal respect, the right power relation, relational approach and transparency, the PhD researcher experienced how these access barriers transformed into value-adding collaborations.
4. Sampling hard-to-reach research communities

Gaining and maintaining access to the research population is the first step in data collection, followed by the selection technique. Sampling further influences the how, who and how many research informants participate in this study (Saunders et al., 2016) and the quality of the data shared to meaningfully answer the research questions (Cunliffe and Alcadipani, 2016; Mikecz, 2012).

Aiming for in-depth understanding of RIEs’ experiences of social inclusion through entrepreneurship in London, by adopting a phenomenological interpretative and intersectional approach, this PhD study aligns its selection strategy to other similar studies, by prioritizing “the selection of subjects who have experience of the phenomenon under investigation” (Englander, 2012:19). Consequently, data saturation is the moot point beyond which any additional data collected would translate into “diminishing returns” for the study (Mason, 2010), emphasised through the researcher’s experiencing of a sense of completeness (Aiken et al., 2015; Nelson, 2016).

Influenced by the research aim and philosophies adopted, the traditional trend in PhD studies reveals an average of ~30 interviews (Mason, 2010), with other qualitative studies including 15-30 participants (Nicolai et al., 2016; Saunders and Townsend, 2016). Following this tradition and considering the following inclusion criteria: country of origin, Romania; immigrant status in the UK; current positionality as business owners/entrepreneurs in London, this study engaged in combining multiple selection techniques, which yielded 51 valid interviews (17 women and 34 men).

With research on hard-to access populations being on the rise (Chorley et al., 2016; Hall et al., 2016), a new publishing stream is taking shape around the methodological insights of accessing and sampling from these communities. These methodological experiences emphasize the benefits of direct, purposive sampling via letters and phone calls (Hall et al., 2015) over indirect sampling, mediated by gatekeepers or community influencers (McFadyen and Rankin, 2016; Quaife et al., 2016). However, for this PhD study, reflectivity in combining sampling techniques rather than unidirectional sampling approach proved critical when researching this hard-to-access population (Hoppitt et al., 2012), as this approach increased the sample size and the quality of the data.

Within the context of migration, Facebook arguably becomes a middle ground between being and becoming, where online relationships are considered as valuable as the real-life ones (Sweet, 2012). Within this context, “social media can provide migration researchers with a unique insight into migrants’ thoughts and behaviours that are occurring naturally in their social networks” (Reips and Buffardi, 2012).

Recently, researchers have started to rely on social media channels to conduct research, either by creating topic-related groups (Brickman and Bhutta, 2012; Valdez et al., 2014) or inviting members of already established groups to participate in a study (Baltar and Brunet, 2012; Crush et al., 2012; Oiarzabal, 2012; Ranieri et al., 2012). However, the e-sampling from the initial research participants’ Facebook accounts implemented by this PhD study is a new research practice (Baltar and Brunet, 2012; Crush et al., 2012), whereby these participants were asked for their consent to sample from their Facebook network, without mediating the initial contact or directly encouraging research collaboration.
This section presents how the PhD researcher reflectively combined the traditional time-space, derived rapport and snowballing techniques with the newly designed e-sampling process via Facebook. This strategy aided the PhD researcher in overcoming the gatekeeper dependency, preventing the constructivist behaviours displayed by members of this community, which could hinder the research progress, increasing the sample size and reducing the selection bias specific to some of the sampling techniques.

Additionally, this PhD study takes on the opportunity of being the first study found to estimate the average time spent in recruiting research participants for each sampling technique used. This is calculated as the estimated logged-in time spent across three main tasks (i.e. gaining research access) and subtasks which defined each of the sampling techniques used (i.e. updating researcher’s Facebook profile). The time was estimated based on time logged for a sample of random informants recruited using each of these four sampling techniques. For each of these randomly chosen informants the researcher logged the time spent on each of the tasks as it happened during the sampling process. This estimated average time does not take into account travelling time, which is rather contextual and could greatly change the estimations made since this PhD study is conducted in London, where travelling time could greatly influence the recruitment time estimates compared to research conducted in smaller cities, for example.

The estimated average time spent per research participant (ATP) is calculated as per the formula shown below, as estimated time spent per informant invited (TPI, in minutes) multiplied by the total number of informants (TNI), then divided by the total number of research participants (TRP):

\[
ATP = \text{TPI} \times \frac{\text{TNI}}{\text{TRP}}
\]

This paper analyses also the research participation rate (RPR) associated with each of the sampling techniques used, which the total number of research participants (TRP) calculates as the total number of informants invited (TII) divided.

\[
RPR = \frac{\text{TII}}{\text{TRP}}
\]

Additionally, the e-sampling via Facebook was developed as an iterative process throughout the sampling period, based on trial and error and on the feedback from e-gatekeepers (Facebook account owners from which the researcher started the e-sampling), who had insights on their network members. Within this sampling process, culturally inspired messages and relational approach were reflectively embedded.

4.1. Sampling by the book

Broadly, accessing research informants is a matter of accessing personal, social, professional and organisational networks (Bryman, 2012), either by cold calling, common to time-space sampling or by participating in community’s events, and even becoming a community member oneself (Ashforth and Reingen, 2014).
Derived rapport sampling (community-based sampling) refers to the sampling led by a gatekeeper, usually an influential community member or leader, who mediates between the researcher and community members (Ellard-Gray et al., 2015).

Usually, researchers using derived rapport to face institutional barriers (Sutherland and Fantasia, 2012), which might delay collaborative actions, even if the organisation could directly benefit from the research (Hoppitt et al., 2012). Additionally, its selection bias limitations are recognized, since the individuals who were more engaged in the community were more likely to be selected and willing to participate in this study, driven by their social openness (Meyer and Wilson, 2009). The researcher’s active membership within this community enabled great support from the community leader, who directly mediated the collaboration with each research
participant, who in turn requested fewer clarifications and were more trustful, compared to those sampled using traditional snowballing, for example. In the context of migration, the church is equally considered a social and a religious hub, bringing together people of different religions, but similar cultural backgrounds.

In practice, the PhD researcher approached the community leader, the priest, from her network of friends to access RIEs affiliated with the Romanian Catholic Church in London. In this case, the priest acted as gatekeeper and mediator for eight out of the twelve research participants initially approached, who further snowballed into eight more. This sampling technique was time efficient, with only 2:48hrs spent on average per research participation.

**Time-space sampling technique**, whereby the researcher targets informants and then identifies venues to approach them in order to invite them to participate in the study (Semaan, 2010).

Although, this sampling technique is known for its high recruitment success rate, particularly in accessing hard-to-access communities (Meyer and Wilson, 2009), for this study it proved to be the most time inefficient. Additionally, to decrease the spatial sampling bias specific to snowball and time-space sampling, the PhD researcher maximized geographic and event variability by participating in heterogeneous networking events (i.e. a cluster of entrepreneurs from different industries and networks) and thus ensured increased sample diversity (Meyer and Wilson, 2009).

By using time-space sampling, the PhD researcher identified three key and influential Romanian entrepreneurs, present in media headlines, and attended events where they were invited as speakers. Given that these events were public, the researcher took the opportunity to approach them directly during the networking sessions. Although very inefficient from a time and budget perspective (on average 18:37hrs spent/ successful research participant), this sampling technique proved useful when reaching targeted individuals, such as “business elites” and “community influencers” who, thanks to their extended immigrant entrepreneurial experience in London, were great sources of research data.

**Snowball sampling** is a non-probabilistic, convenience sampling technique commonly used in qualitative research. Each research participant refers another one or others and so on. This strategy has been successfully used in sampling hard-to-reach populations (Cohen and Arieli, 2011; Tracy, 2013).

Although, theoretically, gatekeepers should be chosen randomly, in practice they are recruited from the researcher’s network of friends and close acquaintances. This translates in sample selection bias, whereby the gatekeepers are likely to “encourage” specific research participants based on personal preferences or their understandings of “being fit” for a particular study or not (Cohen and Arieli, 2011).

Additionally, this sampling technique required spending on average 6:03hrs per research participation, almost three times more than the average time spent when using e-sampling via Facebook.
4.2. E-sampling: Facebook to the rescue

Facebook is a social media network, with 2.7 billion worldwide monthly users, each of whom spends daily, on average, 25 minutes (Statista, 2019); creating, maintaining and sharing varied forms of information (Mintel, 2013; Xiang and Gretzel, 2010). This seems to shape the new network society (Castells et al. 2007) for over 30% of the world population, where each of these users has, on average, a network of 338 “friends” (McClain, 2017). These statistics put Facebook into the top preferred social media for individuals (Statista, 2019).

Technological advances change how this globalised world interconnects and communicates and they encourage research output and multidisciplinary research (González-Bailón et al. 2014). This new-networked society enables the rise of new research opportunities, where viable online research alternatives, including e-sampling and online/messenger interviews enable greater research resource efficiencies compared to the traditional ones. Facebook’s strength in becoming a great data source and recruiting pool for researchers (Harris et al., 2015; Young et al., 2014) is in its attributes, including size (i.e. 30% of the world’s population), influence (i.e. undirected reach), sharing (i.e. real time population reach). It is increasingly becoming an effective and efficient data source for research (Ferrara et al. 2014; Kurka et al. 2015).

Additionally, Facebook “society” closely mirrors the offline demographical profile of many societies (Duggan, 2015). The Facebook groups and individual accounts create for their members a real time opportunity for virtual social interaction, which overcomes time, space and even the social power constraints of traditional society (Bennett and Segerberg, 2012; van Dijck, 2013).

Additionally, with the increasing influence of social media, including Facebook, within the context of migration (Dekker and Engbersen, 2014; Oiarzabal, 2012) comes the opportunity of online accessing and sampling hard-to-access population, which could define time effective research solutions (Rife et al., 2016; Schneider et al., 2015). A more creative and reflective research approach towards social media could enable a more inclusive study, allowing hard-to-reach communities, such as the immigrants, to be reached and heard (Crush et al., 2012; Kayrouz et al., 2016; Oiarzabal and Reips, 2012).

Most research uses social media to conduct surveys (Baltar and Brunet, 2012; Crush et al., 2012; Valdez et al., 2014), mainly in medical and healthcare (Arcia, 2014; Head et al., 2016; Lohse, 2013) and cross-cultural fields (Thomson and Ito 2014). Within this context, the research opportunity of using Facebook as a venue to sample for face-to-face interviews remains largely underexplored, but necessary (Piña-García et al., 2016).

Thanks to Facebook’s mission of “giving people the power to share and making the world more open and connected” (Facebook, 2019), it becomes an unlimited source of data to support knowledge creation and sharing for many stakeholders (Lu and Brelsford 2014; Piña-García and Gu 2013; 2015), including researchers, as long as innovative methodologies are timely and coherently formulated and ethically applied.

Against this landscape, e-sampling via Facebook addresses some of the challenges presented by gatekeepers by transferring some of that power back to the researcher and by enabling the increase of the sample size, which are particularly relevant when researching “hard-to-access” populations (Ellard-Gray et. al. 2015).
E-sampling via Facebook is a sampling technique whereby the researcher screens the Facebook networks of “befriended” research participants to identify new members of the target population, then reaches out to inform and invite them to participate in the study by using Messenger messages.

Compared to the traditional snowballing, in e-sampling via Facebook the research participant’s role is passive, since no direct referral is expected. The researcher accesses the research participant’s network, which is publicly available, screens it for potential informants based on a set of variables, which are likely to be listed in her/his Facebook profile or online posts, to create a short list of suitable research participants, which are then invited to participate in the study through a standardised, Private Message via Messenger. Once these potential informants consent to participate, their network is used as sampling source for others.

The PhD researcher designed and implemented the following e-sampling process via Facebook:

1. Updating the researcher’s Facebook profile with academic association to the University and prior professional career in Investment Banking to increase credibility and trust in researcher and research;
2. Updating researcher’s profile photos to mirror the broad profile of potential research participants (i.e. professional dress code, shared links on entrepreneurship research) and thus increasing the odds of researcher-participant relating to each other;
3. “Befriending” the research participants on Facebook in order to maintain the contact for future collaboration and to enable access to their network for sampling opportunities; avoiding public categorisation of research participants in order to maintain research anonymity.
4. Identifying the research participants and professional network affiliations (business or entrepreneurial networks) based on the inclusion criterion detailed above.
5. Asking permission from these research participants to approach “friends” from their network for research access purposes;
6. Identifying potential research participants with the following attributes:
   - Romanian nationality - Facebook account listing a Romanian name and postings in the Romanian language
   - Immigrant status - London, UK listed as current residence, combined with previous residence or studies completed in Romania
   - Entrepreneur – company name and position in the company listed on the account (i.e. “founder/owner/entrepreneur” at X company in London, UK) and business adverts.
7. Validating their active business account using www.companycheck.co.uk;
8. Writing to them on Messenger a standard, culturally sensitive and empathetic brief to introduce the researcher, the research and to invite them to a more detailed conversation around the opportunity to participate in this study;
“Dear Mrs. / Mr. X,

I apologize for addressing you via Facebook and not in a more professional manner, via email.
My name is X and as a PhD researcher at the University X, I would be extremely pleased to include your story as a Romanian entrepreneur in my PhD study. I am convinced that your story can influence not only the image of the Romanian community in the UK but also inspire other Romanians.
Please do not hesitate to ask any questions. I would be privileged to meet you and hear your story.
Only an hour of your life can influence other lives and opinions, or simply help another Romanian.
Sorry if I dared. Have a nice week! I would appreciate an answer.
Sincerely, X, a great Romanian like all the other Romanians.”

Fig. 1. Sample of the standardised message (Translation from the original Romanian message)

9. Waiting for them to open the dialogue, either by requesting an introductory phone conversation or by requesting more details regarding the research (on average a week’s waiting time);
10. Encourage trust by “befriending” them to increase the researcher’s profile visibility and communication opportunity and familiarity;
11. E-mailing a formal invitation, including all the ethics documentation approved by the University’s Ethics committee, together with alternative interview opportunities;
12. Scheduling the interview and e-mailing a formal confirmation;
13. Sending weekly Messenger well-being messages to build trust and maintain access;
14. Encouraging personal and professional discussions as an opportunity to establish the power relation rapport they prefer and to put them at ease to share information;
15. Sending out a “thank you” message within 24 hrs after the face-to-face interview, as token of appreciation for their research participation;
16. Maintaining the professional relationship by sending weekly Messenger well-being messages to ensure the nurturing of further research collaboration.

By acknowledging the criticality of the recruitment message (Teo et al., 2018; Choi et al., 2017) and how culturally and context-sensitive this is (Batterham, 2014; Burgess et al., 2017), the wording in the standardised message used in e-sampling via Facebook emphasises the researcher’s positionality as a cultural and linguistic insider in relation to the informants’ community.
The participants’ act of generosity could be motivated by the desire to help or to initiate positive change. However, being fully aware of the importance of hierarchical power relations in Romanian culture, the PhD researcher used key cultural words to emphasise that the decision belongs purely to them, the informants and that their own power position, if desired, is ensured during this collaboration. Additionally, it was very important to reiterate the fact that the researcher is a cultural insider and that participant-centrality is prioritized, starting with respecting their decision of participating or not and the location, time and interview set up.

This strategy is known as a participatory model, which encourages “non-manipulative research relationships which have the potential to overcome researcher-researched separation” (Reinharz, 1983, cited in Dwyer and Buckle, 2009:62).

Within the context of this study, the virtual snowballing sampling techniques was critical, firstly because the traditional ones dried up, putting the research on hold for two weeks, and secondly because it was the most time efficient compared to all the other traditional sampling techniques used with an estimated average time spent of 2:31hrs per research participant.

Although there are concerns about the selection bias associated with online sampling because it only accesses a segment of the population which is present online and thus not the whole target population (Ruths and Jurgen, 2014), it can be argued that selection bias could be reduced when sampling on Facebook. This is possible thanks to the large sample frame, the opportunity of accessing diverse, multiple networks chosen directly by the researcher, and of focusing the research sampling requirements based on gender, industry, age, profession or other relevant variables (Singh and Wassenaar, 2016).

Additionally, some researchers argue that the user might tailor the self-reported Facebook profile to suit his social interests (Kosinski et al., 2015), whilst empirical evidence supports the contrary, showing that Facebook profiles are trustworthy accounts (Back et al., 2010; Kosinski et al. 2013). The PhD researcher’s experience of sampling via Facebook was a positive one, as all the Facebook profiles of the RIEs were trustworthy and up-to-date in areas of the researchers’ interest.

This reflective approach allowed active management of some of the research bottlenecks associated with traditional sampling, such as selection bias and it encouraged the design of time efficient e-sampling solution via Facebook. In this case, the e-sampling via Facebook improved the research findings, by improving the selection from an industry, size and gender perspective (Loxton et al., 2015), and data quality, as this selection technique promoted the inclusion of more experienced and long term entrepreneurs in this study. This new sampling technique moreover proved to be the most time efficient one of these attempted.

4.3. Analysis of sampling time efficiencies

This paper emphasises the benefits of reflectively combining different sampling techniques in order to increase the research sample, by maximizing the opportunities of reaching these hard-to-reach populations, particularly when scarce resources, developing skills and tight deadlines are realities of many research projects, including this PhD study (Bovevski et al., 2014; Shedlin et al., 2011).
This analysis considers jointly time efficiencies and the research participation rate. This is done per research participant and sampling task for each of the sampling techniques used in this PhD study.

The time efficiencies experienced by the qualitative PhD researcher stem from reflectively combining three traditional sampling techniques with a bespoke, virtual sampling one via Facebook. Detailed views of the participation rate and time efficiencies experienced with each of the four sampling techniques are detailed in the tables and graph below (Appendix 1. Detailed summary).

Table 1. Participation rate per sampling technique

<table>
<thead>
<tr>
<th>Sampling Technique</th>
<th>No. of informants invited</th>
<th>No. of research participants</th>
<th>Research participation rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional snowballing</td>
<td>94</td>
<td>29</td>
<td>31%</td>
</tr>
<tr>
<td>Derived rapport</td>
<td>12</td>
<td>8</td>
<td>67%</td>
</tr>
<tr>
<td>E-sampling via Facebook</td>
<td>47</td>
<td>13</td>
<td>28%</td>
</tr>
<tr>
<td>Time/space sampling</td>
<td>11</td>
<td>3</td>
<td>27%</td>
</tr>
<tr>
<td>Total population</td>
<td>164</td>
<td>53</td>
<td>32%</td>
</tr>
</tbody>
</table>

Source: PhD researcher’s fieldwork

The participation rate has different meanings across studies, but it is often discussed for its impact on research quality and validity (Gales and Tracy, 2007, cited in Morton et al., 2012). In recent decades, the increased attention to data confidentiality and ethics standards have redefined the research arena reducing participation rates from 60-70% to 20% and thus increasing the risk of selection bias (which need to be managed) rather than the selection bias (Visser et al, 1996, cited in Morton et al., 2012). However, evidence suggests that studies with lower response rate from 5% to 54% are as valid as those with higher response rates, although they are more exposed to this risk (Holbrook et al, 2007) since participation rate is informative not normative (Scott et al., 2006). Additionally, given that this PhD study aims to interpret how social inclusion is experienced by RIEs in London, the discussion of limitations should raise questions of achieving the intended depth and not generalisation in this case (Guetterman, 2015; Moser and Korstjens, 2018). Within this context, the average response rate of 32% is in line with previous research on hard-to-reach populations (see Topolovec-Vranic and Natarajan, 2016).

From a research participation rate detailed in Table 1 above, as expected, by using the derived rapport through the church community, the PhD researcher experienced the highest research
participation rate (67%) out of all the sampling techniques used, for the reasons already discussed. By comparison, all the other sampling techniques reported similar research participation rates between 27% and 31%. The traditional snowball sampling yielded 31% participation rate, which is similar to other studies (Kirchherr and Charles, 2018). The esnowball sampling technique yielded 28% participation rate, which is also comparable to other studies (Baltar and Brunet, 2012; Kirchherr and Charles, 2018).

Considering that 79% (42 research participants out of overall 53 interviewed) of the research participants were recruited using snowballing and e-sampling combined, the previous empirical evidence is confirmed, whereby, snowballing, despite all its limitations, is feasible and ideal particularly when researching hard-to-reach populations like the immigrants (Faugier and Sargent, 1997, cited in Ellard-Gray et al., 2015). However, see below for the highly favourable, lower time investment in Facebook sampling.

As a contribution to knowledge, this PhD study captures the time efficiencies associated with each of the four sampling techniques used. This is critical in research planning and it has not be found in any other qualitative study, although it might just be the most time- and energyconsuming part of the research process. Only a handful of studies, quantitative in nature, present the financial costs associated with sampling, thanks to a social media algorithm that captures this recruitment, using the internet “click” on adverts, for example (Potzschke and Brown, 2017).

Based on the method detailed above, this PhD researcher calculated the estimated average time spent per each of the three tasks defining the overall sampling process using the following formula:

\[ ATS = ATG + ATN + ATM \]

<table>
<thead>
<tr>
<th>Sampling techniques</th>
<th>Gaining research access ATG</th>
<th>Negotiating research participation ATN</th>
<th>Maintaining research access ATM</th>
<th>TOTAL (est. avg. time) ATS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional snowballing</td>
<td>0:39:50</td>
<td>4:38:06</td>
<td>0:45:21</td>
<td>6:03:17</td>
</tr>
<tr>
<td>Derived rapport</td>
<td>0:26:15</td>
<td>2:01:15</td>
<td>0:20:00</td>
<td>2:47:30</td>
</tr>
<tr>
<td>E-sampling via Facebook</td>
<td>1:34:32</td>
<td>0:32:28</td>
<td>0:23:32</td>
<td>2:30:32</td>
</tr>
<tr>
<td>Time-space sampling</td>
<td>17:20:00</td>
<td>1:06:40</td>
<td>0:10:00</td>
<td>18:36:40</td>
</tr>
<tr>
<td>Total population</td>
<td>1:47:49</td>
<td>3:02:12</td>
<td>0:34:10</td>
<td>5:24:11</td>
</tr>
</tbody>
</table>
From a time efficiency perspective, the e-sampling via Facebook proved to be the most time-efficient out of all four sampling techniques used, requiring only an estimated average time spent of 2:31 hrs per research participant. This time efficiency was driven mainly by a very short time required to negotiate research participation, a ninth of the negotiating time required for the traditional sampling, because there was no need to negotiate with the gatekeepers, outside the initial consent to use their network. The response to the initial online message sent by the researcher was certainly based on screening the researcher’s Facebook profile as well as the association with the informant’s “friend” as evidenced by Facebook analytics. This approach encouraged early research collaboration, whereby the time dedicated by the researcher on personal and research introductions was replaced by the informants spending time on these tasks instead, using the information readily available on the researcher’s Facebook profile. The trust seems to build faster, thanks to the abundant personal information on Facebook, replacing the introductory phone calls, which became a time-consuming routine during traditional snowball sampling.
The snowball sampling technique required an estimated average time spent of 6:03 hrs per research participant, almost three times more than the estimated average time spent when using e-sampling via Facebook. This time inefficiency was driven by doubling the negotiation time and effort with gatekeepers and the referred informants.

The derived rapport sampling technique was time efficient, as expected, with only 2:48hrs average estimated time spent per research participant. This efficiency was driven by the fact that the researcher is an active member of the church community and because the priest is a valued leader in the catholic communities in general. Consequently, the priest’s invitation to church members to participate in this research was prioritized. One could argue that this participation rate is skewed by a small sample framework of only 12 RIEs out of which 8 participated in this study, yielding a 67% participation rate.

The time-space sampling was very time-inefficient, requiring an estimated average time spent of 18:27 hrs per research participant. This time inefficiency was driven by attempts to gain access due to many events, which themselves had disappointingly low or zero participation rates. The rest of the time-space sampling tasks were most efficient by comparison with the other sampling techniques, particularly because the participants recruited tended to lack time for off-topic conversation.

5. Conclusions and future research recommendations

Undeniably, collecting good quality primary data could become a complex and challenging process (Singh and Wassenaar, 2016), particularly for less experienced researchers and those labelled as “outsiders” to research participant communities (Shedlin et al., 2011). This becomes particularly true when the researcher has to negotiate access with gatekeepers to “hard-to-access” communities, such as the immigrants (Vuban and Eta, 2018), or when the traditional sampling techniques dry up (Grigori and Baltar, 2013). However, this experience creates the opportunity of reflecting upon the unlimited potential of knowledge creation though newer research tools, such as Facebook, which seemingly mirrors a borderless world, where reality becomes more fluid and communication becomes ubiquitous.

The PhD researcher recommends approaching research as a dynamic, iterative and reflective process (Hoppitt et al., 2012), by finding context-bounded access strategies to overcome barriers, by using multiple sampling techniques to improve participant selection (Bonevski et al., 2014; Ellard-Gray et al., 2015).

This study emphasises the importance of approaching research selection as an ongoing process of engagement, rather than a one-time event; a relation-based journey, rather than an instrumental one. E-sampling via Facebook proved critical in overcoming the bottleneck of the traditional sampling techniques. Thanks to its reduced response time and less negotiation, it was also the most time efficient participant selection tool with only 2:31 hrs spent on average per research participant.

In future research, reflection regarding all stages of qualitative research should be considered as real opportunities to increase research quality and inclusiveness (Rockliffe et al., 2018).
Whilst researching hard-to-reach communities and engaging with new research practices can be promising, by maximizing selection efforts, and challenging, due to lack of up-to-date guidance in use of participant selection tools and the ongoing debate about data privacy. Researchers should explore new social environments, such as Facebook and other social media platforms, as these mirror a seemingly borderless world, full of untapped research opportunities at a larger scale, delivering more time efficiencies than ever before. There is need for more studies and best practices to emerge and thus to fully explore the research potential of these social platforms, particularly when it comes to participants’ demographics and time effectiveness of online selection strategies (Topolovec-Vranic and Natarajan, 2016).
## Appendix

### Appendix 1. Detailed summary of the estimated average time spent per research participant for each sampling technique

<table>
<thead>
<tr>
<th>Sampling Technique</th>
<th>Total no. of potential informants reached</th>
<th>Time spent on average</th>
<th>Hours spent on average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional snowballing</td>
<td>94</td>
<td>29</td>
<td>175:35:00</td>
</tr>
<tr>
<td>Identifying community gatekeepers</td>
<td>1</td>
<td>06:00</td>
<td>6:00:00</td>
</tr>
<tr>
<td>Reaching out to community gatekeep</td>
<td>13</td>
<td>00:15</td>
<td>3:15:00</td>
</tr>
<tr>
<td>Briefing on the research and clarifica</td>
<td>4</td>
<td>02:30</td>
<td>10:00:00</td>
</tr>
<tr>
<td>Negotiating participation and clarifica</td>
<td>4</td>
<td>01:30</td>
<td>6:00:00</td>
</tr>
<tr>
<td>Contacting the referred participants</td>
<td>94</td>
<td>00:40</td>
<td>62:40:00</td>
</tr>
<tr>
<td>Briefing on the research and clarifica</td>
<td>38</td>
<td>01:40</td>
<td>63:20:00</td>
</tr>
<tr>
<td>Email confirming research participation</td>
<td>29</td>
<td>00:05</td>
<td>2:25:00</td>
</tr>
<tr>
<td>Well-being ongoing messaging, phone</td>
<td>35</td>
<td>00:01</td>
<td>0:35:00</td>
</tr>
<tr>
<td>Support participants’ business project</td>
<td>16</td>
<td>01:20</td>
<td>21:20:00</td>
</tr>
<tr>
<td><strong>Average time per participant</strong></td>
<td><strong>29</strong></td>
<td></td>
<td><strong>6:02:17</strong></td>
</tr>
<tr>
<td>Derived Rapport</td>
<td>12</td>
<td>8</td>
<td>22:20:00</td>
</tr>
<tr>
<td>Reaching out to the community</td>
<td>1</td>
<td>01:00</td>
<td>1:00:00</td>
</tr>
<tr>
<td>Briefing on the research and clarifica</td>
<td>1</td>
<td>02:30</td>
<td>2:30:00</td>
</tr>
<tr>
<td>Negotiating participation and clarifica</td>
<td>1</td>
<td>01:30</td>
<td>1:30:00</td>
</tr>
<tr>
<td>Contacting the referred participants</td>
<td>12</td>
<td>00:30</td>
<td>6:00:00</td>
</tr>
<tr>
<td>Briefing on the research and clarifica</td>
<td>8</td>
<td>01:00</td>
<td>8:00:00</td>
</tr>
<tr>
<td>Email confirming research participation</td>
<td>8</td>
<td>00:05</td>
<td>0:40:00</td>
</tr>
<tr>
<td>Well-being ongoing messaging, phone</td>
<td>10</td>
<td>00:01</td>
<td>0:10:00</td>
</tr>
<tr>
<td>Support participants’ business project</td>
<td>2</td>
<td>01:15</td>
<td>2:30:00</td>
</tr>
<tr>
<td><strong>Average time per participant</strong></td>
<td><strong>8</strong></td>
<td></td>
<td><strong>2:47:30</strong></td>
</tr>
<tr>
<td>E-snowballing via Facebook</td>
<td>47</td>
<td>13</td>
<td>32:37:00</td>
</tr>
<tr>
<td>Updating own profile</td>
<td>1</td>
<td>00:30</td>
<td>0:30:00</td>
</tr>
<tr>
<td>Identifying entrepreneurs in own network and asking permission to</td>
<td>1</td>
<td>02:00</td>
<td>2:00:00</td>
</tr>
<tr>
<td>Activity Description</td>
<td>Number</td>
<td>Time</td>
<td>Total Population</td>
</tr>
<tr>
<td>-----------------------------------------------------------</td>
<td>--------</td>
<td>-------</td>
<td>------------------</td>
</tr>
<tr>
<td>Screening networks for potential informants</td>
<td>30</td>
<td>00:25</td>
<td>164</td>
</tr>
<tr>
<td>Reaching out to potential informants</td>
<td>47</td>
<td>00:07</td>
<td>53</td>
</tr>
<tr>
<td>Clarifying and negotiating participation</td>
<td>21</td>
<td>00:17</td>
<td></td>
</tr>
<tr>
<td>Email confirming research participation</td>
<td>13</td>
<td>00:05</td>
<td></td>
</tr>
<tr>
<td>Well-being ongoing messaging</td>
<td>21</td>
<td>00:06</td>
<td></td>
</tr>
<tr>
<td>Support participants’ business project</td>
<td>3</td>
<td>01:00</td>
<td></td>
</tr>
<tr>
<td><strong>Average time per participant</strong></td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Time-space sampling</strong></td>
<td>11</td>
<td>03:00</td>
<td></td>
</tr>
<tr>
<td>Identifying the community role model</td>
<td>1</td>
<td>09:00</td>
<td></td>
</tr>
<tr>
<td>Identifying &amp; signing up for the relevant events</td>
<td>1</td>
<td>04:00</td>
<td></td>
</tr>
<tr>
<td>Participating to relevant events</td>
<td>13</td>
<td>03:00</td>
<td></td>
</tr>
<tr>
<td>Meeting &amp; briefing on the research</td>
<td>5</td>
<td>00:10</td>
<td></td>
</tr>
<tr>
<td>Negotiating participation</td>
<td>3</td>
<td>00:30</td>
<td></td>
</tr>
<tr>
<td>Customize email confirming research</td>
<td>3</td>
<td>00:20</td>
<td></td>
</tr>
<tr>
<td>Well-being ongoing messaging</td>
<td>3</td>
<td>00:10</td>
<td></td>
</tr>
<tr>
<td><strong>Average time per participant</strong></td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total population</strong></td>
<td>164</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td><strong>Average time per participant</strong></td>
<td>53</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: PhD researcher’s fieldwork
References


• http://dx.doi.org/10.1037/a0039210.


