

## Qualitative Experiments for Social Sciences

Nadia Steils <sup>1</sup> <https://orcid.org/0000-0002-9239-3635>

<sup>1</sup> IAE Lille, University of Lille, Lille Economy and Management, France.

### Abstract:

This paper presents the qualitative experiment as an alternative methodological solution that combines an open qualitative approach, and a structured and controlled experiment. Using three studies, including both a qualitative experiment and a traditional in- depth interviews approach, we compare the findings of both approaches to identify the benefits and risks of qualitative experiments. Our findings contribute by presenting a methodological framework and technical recommendations based on three validity criteria (internal, external, and interpretivist validity). The results thereby contribute methodologically by empirically investigating the usefulness of qualitative experiments based on a combination of quantitative and qualitative validity criteria identified in the literature.

**Keywords:** qualitative experiments, validity criteria, methodology, qualitative research.

**Submission:** 25/10/2020

**Acceptance:** 03/01/2021

## 1. Introduction

While qualitative and quantitative techniques are often considered separate approaches, the combination of different research designs and collection processes have been recommended to discover differentiated results. According to previous authors, this is beneficial to the holistic investigation of complex studies (Decuir-Gunby & Schutz, 2017). In line with previous authors' recommendation to continue exploring mixed research approaches (Ramirez-Montoya & Lugo-Ocando, 2020), this paper presents the qualitative experiment as an alternative methodological solution that combines an open qualitative approach, and a structured and controlled experiment (Robinson & Mendelson, 2012).

More precisely, according to quantitative research, a true experimental design is one that has at least two independent, parallel groups; randomly assigns subjects to the

groups; and assesses treatments prospectively (Slack & Draugalis, 2001). Applied to qualitative research, the investigation and analysis is supported by an open and interpretivist approach (Robinson & Mendelson, 2012). The combination of a qualitative and experimental approach seems especially useful in the investigation of processes and behaviors in social sciences (Kleining & Witt, 2000; Wagoner, 2015).

Consequently, this research paper presents a methodological framework for conducting qualitative experiments, including recommendations and techniques to address internal, external and interpretivist validity. Based on three studies that apply the three types of validity and compare the insights gained from the qualitative experiments with those from the traditional in-depth interview approach, the findings also identify the benefits and risks of qualitative experiments for interpretivist research. This paper is structured as follows. After reviewing the literature on qualitative experiments, validity criteria commonly used in qualitative and experimental (quantitative) research are examined to develop a methodological framework that is applied to the three studies of this research. Finally, the systematic methodology is presented before discussing the findings of this research.

## **2. Theoretical background**

The qualitative experiment is the intervention in a social/psychological situation for exploratory purposes. When compared with experiments, qualitative experiments are exploratory and heuristic (Kleining, 1986). They are based on the inter-related strategies of “experiments” and “observation” that improve learning and understanding (Kolb, Boyatzis & Mainemelis, 2002; Kleining, 1991). They differ from systematic observations in that the analysis remains discovery-based and introspective (Kleining & Witt, 2000), and differ from traditional qualitative methods in that they allow a controlled and systematic study of patterns (Robinson & Mendelson, 2012). They are especially suitable to discover qualitative relations such as structures, processes or structural changes (Kleining & Witt, 2000; Wagoner, 2015). For example, Bartlett classified participants based on a priori characteristics using qualitative experiments (Bartlett, 1995).

While the term has been coined by three main authors before (Bartlett, 1953, 1995; Kleining, 1986; Robinson & Mendelson, 2012), we argue that qualitative experiments can benefit from validity criteria and techniques used in quantitative research to increase the results’ relevancy of qualitative experiments. Kleining and Witt (2000) identified four rules for the implementation of qualitative experiments:

- Researchers’ openness regarding their preconceptions if the data are not in agreement with them;
- Changing research questions during the research process;
- Rich and varied data based on a variation in sample and method to avoid one-sided representations;
- Analysis oriented towards the discovery of similarities.

## **3. Qualitative and Quantitative Validity Criteria and Techniques**

Quantitative research has long used validity criteria to evaluate the relevance of

statistical findings. While quantitative research mostly used internal and external validity criteria, qualitative researchers adapted validity criteria and presented different norms and techniques to achieve it. Table 1 summarizes the main validity criteria in quantitative and qualitative research, and the techniques used to increase each type of validity as described below. On the one hand, in quantitative research, internal validity refers to the degree to which a study establishes the cause-and-effect relationship between the treatment and the observed outcome. In other words, it refers to the process of ruling out alternative causes (Slack & Draugalis, 2001). Internal validity, as defined by Campbell and Stanley, is a logical rather than statistical issue. Researchers are encouraged to assess internal validity based on the research design and operational procedures. External validity is addressed by delineating inclusion and exclusion criteria to characterize participants and assess generalizability of the results (Slack & Draugalis, 2001; Lynch, 1983).

On the other hand, in qualitative research, validity is characterized by a tension between rigor and creative discovery (Whittemore, Chase, & Mandle, 2001). Primary validity criteria are based on credibility, authenticity, integrity and critique, while secondary validity criteria focus on congruence, creativity, vividness, explicitness, and thoroughness (Whittemore, Chase, & Mandle, 2001). Credibility and authenticity refer to the descriptive and interpretive validity evaluation of qualitative research (Maxwell, 1996). Lincoln (1995) describes credible findings as those reflecting the relativistic nature of truth claims in the interpretivist tradition (Lincoln, 1995). Researchers should make sure that the findings reflect the experience of participants or the context in a believable way (Lincoln, 1995). The findings should also be authentic in their representation and reflect the meanings and experiences that are perceived and lived by the participants of the study (Sandelowski, 1986).

To assess these different types of validity, researchers recommend a series of techniques that contribute to valid research findings. Qualitative and quantitative research both claim that validity is mainly assessed through the choice of research design (Whittemore, Chase, & Mandle, 2001). Qualitative research insists also on triangulating methods or validity checking (member or experts checks) (Whittemore, Chase, & Mandle, 2001). Qualitative research thus cannot be assessed through a single test or step in the research process, but should rather follow the rules of a processual validity that depends on decisions made at every step of the research process (Hayashi, Abib, and Hoppen, 2019). An attempt to remain true to the phenomenon under study is essential in qualitative research (Hammersley, 1992), which suggests the integration of principles commonly used in quantitative research for the qualitative study of processes, patterns and behaviours.

Quantitative, namely experimental research, insists on sample-related decisions to address external validity (Leviton, 2017). Experimental research recommends including random or systematic sampling across two or more study conditions, to capture the most relevant variation (Leviton, 2017). More precisely, a true experimental design is one that has at least two independent, parallel groups; randomly assigns subjects to the groups; and assesses treatments prospectively (Robinson & Mendelson, 2012; Seltman, 2018; Slack & Draugalis, 2001). In within-subjects experiments the randomization of treatments is recommended (Dülmer, 2015). Moreover, external validity requires that

researchers encourage heterogeneity of study settings or populations to compare the outcomes across study contexts (Leviton, 2017). Beyond surface variables such as demographics, researchers should also consider other features (e.g., privacy concern) to study individuals, as they may interact and alter consumers’ perception or behaviours (Leviton, 2017; Seltman, 2018).

Lynch (1982) argues that for external validity to increase, there should be statistical generalizability that can be induced through appropriate use of sampling procedures (Ferber, 1977); conceptual replicability or robustness when using particular subjects or settings (Cook and Campbell 1979); and realism of tasks, stimuli and settings (Berkowitz and Donnerstein 1982). As external validity is hard to assess, even in quantitative research, reweighting the sample to match the population has been used as a solution to increase external validity (Andrews & Oster, 2018). The latter design elements also contribute to greater internal validity. Controlling for potentially confounding variables minimizes the potential for an alternative explanation for treatment effects and provides more confidence that effects are due to the independent variable (Slack & Draugalis, 2001).

Regarding the analysis of qualitative experiments sampling is determinant for validity. In particular, new variations of data and perspectives should not bring new results but confirm the existing ones (saturation principle). Findings are reliable if all data can be imputed to the same categories (Kleining, 1986). Reviewing validity criteria in qualitative and experimental quantitative research helped to identify the potential criteria that need to be considered when conducting scientifically sound qualitative experiments. Table 1 synthesizes validity criteria in both approaches, as well as related recommendations to address them. These criteria and techniques set a methodological framework for conducting qualitative experiments, which we will apply to three studies with the goal of identifying the benefits and risks of using qualitative experiments.

**Table 1:** Validity criteria in qualitative and experimental quantitative research.

	Qualitative research	Quantitative (experimental) research
	Interpretivist validity	Internal validity
Research design and analysis	<ul style="list-style-type: none"> <li>• Open questions</li> <li>• Open coding/grounded analysis</li> </ul>	<ul style="list-style-type: none"> <li>• Adherence to a study protocol</li> <li>• Between-subjects (vs. within-subjects design)</li> <li>• Reducing the impact of extraneous factors</li> </ul>
	Interpretivist validity	External validity
Sampling	<ul style="list-style-type: none"> <li>• Sampling diversity</li> <li>• Saturation principle</li> </ul>	<ul style="list-style-type: none"> <li>• Systematic sampling procedure / randomization</li> <li>• Realism of tasks, stimuli and setting (heterogeneity of settings)</li> </ul>

#### 4. Goals and Methods

This paper aims to examine the benefits and risks of qualitative experiments, and provide methodological recommendations based on qualitative and experimental validity criteria. Based on a quantitative and qualitative literature review, three types of validity criteria have been identified (interpretivist, internal and external), which can be addressed through nine methodological techniques (see Table 1). Qualitative experiments have shown to be useful for examining processes and behaviours (Kleining,

1986). Therefore, we decided to compare the usefulness of qualitative experiments to investigate major processes and behaviours in consumer research (information search & learning, and decision-making), as these research streams are particularly and increasingly interested in investigating consumer and firm co-created activities and behaviours (Cho, Fu & Wu, 2017). To compare between- and within subjects experiments (Dülmer, 2015; Slack & Draugalis, 2001), we studied decision-making using one between- and one within-subjects experiment, leading to three parallel qualitative studies (see Table 2).

Each of the three qualitative experiments (N1=16; N2= 15; N3= 15) used a traditional qualitative approach (in-depth interviews) preceded by a qualitative experiment. First, we used a thematic analysis (open and axial coding) for the analysis of each approach independently. Next, we compared the findings from the in-depth interviews and the qualitative experiment to derive the benefits and risks of each approach. Study 1 aimed to understand consumers' information and learning behaviours online when having to learn the usage of a new product. Therefore, three situations have been imagined, in which consumers had to learn and carry out a task they had never done before (installing wooden parquet flooring, using a SodaStream maker, or editing photos on Photoshop). These situations were chosen to ensure setting diversity and realism (Leviton, 2017), and thus internal validity as described in the literature. To address internal validity even more, participants conducted one of the three tasks (between-subjects; Robinson & Mendelson, 2012), in the same artificially decorated study lab (reducing the impact of extraneous factors), and based on a study protocol in which they had one hour to conduct the task while being filmed and observed. External validity was increased by providing consumers with all informational material that they could have found at home as well (not only a computer, even though the study goal was to investigate online learning). We provided them with real tasks and situations (increasing realism, Berkowitz and Donnerstein 1982). Furthermore, we systematically assigned them to one or the other task based on the only criteria of having not used the product before ("systematic" sampling, Leviton, 2017). Finally, interpretivist validity was ensured by recruiting a diverse study sample based on the saturation principle, providing few initial instructions on how to behave during the experiment, and openly discuss this and similar experiences in follow up semi-structured interviews (Lincoln, 1995).

Study 2 and 3 investigated consumers' purchase decision-making after watching influencer videos. Study 2 invited participants in one of two situations using both similar videos, but from influencers with varying community size (control for alternative explanations, increasing internal validity; Seltman, 2018). Study 2 aimed the same research objective, but showed each participant the three videos with similar content, but different community sizes (within-subjects experiment; Dülmer, 2015). After watching the videos, consumers were asked to shop online for the necessary material using the "think-aloud" technique, but without providing any further instructions to ensure an open research approach, and thus interpretivist validity (Maxwell, 1996). To further increase interpretivist validity, both studies conducted complementary in-depth interviews on this and similar experiences.

## **5. Analysis and Results**

In order to derive the risks and benefits of experiments for qualitative research, we coded and analyzed the results of each approach (qualitative experiment vs. in-depth interviews separately). In particular, the researcher transcribed video and audio material and coded decision-making and behaviors according to the research objective(s) using a thematic analysis. Next, we compared the findings and identified similarities and differences (see Table 2).

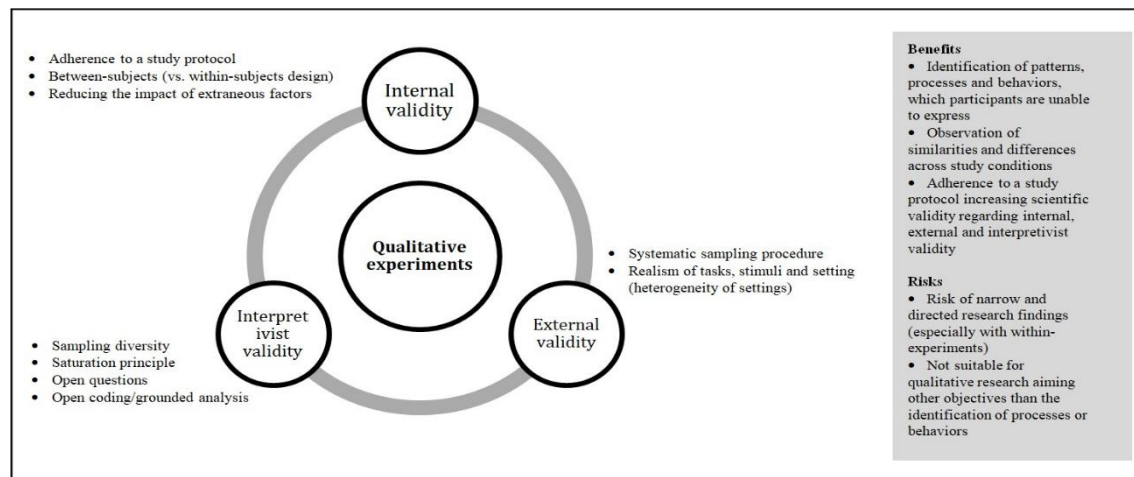
The results from study 1 identified how individuals learn online in consumption-related contexts. During in-depth discussions, participants explained their information search in a seemingly well- thought processing order using dominant preferences for learning tools and sources. Some participants were also unable to explain their process and preferences. However, the experimental conditions revealed an iterative search behavior using a variety of information sources. Comparing the three scenarios, we further observed similarities of searching behaviors across scenarios (e.g., scant reading, jumping video sections), as well as divergences of searching behaviors within scenarios (e.g., searching order). To sum up, we observe that qualitative experiments helped us to identify processes and behaviors, which participants were unable to express. Furthermore, it helped to explore two main learning processes (driven by a learnability or a usability goal) by identifying similarities and differences across the three study conditions.

Study 2 and 3 investigated consumers' decision-making process in social media purchases (influencer videos) using a between-subjects (study 2) or within-subjects (study 3) qualitative experiment. Given the realistic setting of the experiment, we observed that participants in study 2 "freely" searched for and evaluated products on the web-based on (or not) the influencer videos. None of them followed the recommendations from the video, but rather searched for products on webshops they were used to. However, in the in-depth discussions, a part of the consumers mentioned that the products recommended in the video seemed interesting and participants even considered buying them. Furthermore, we observed similarities and differences between the conditions. For example, a longer decision-making process in the qualitative experiments, which, as expressed by participants, seemed shorter when explained. The same benefits of qualitative experiments observed in study 1 can be concluded in study 2. However, we observed several risks and inconveniences in study 3, which used a within-subjects design. Here, we observed that presenting different videos with slightly changing settings encouraged participants to focus their attention and discussion on these elements, thereby reducing the overall openness of the qualitative approach (reduction of the interpretivist validity). Detailed observations and findings are summarized in Table 2.

**Table 2:** Study overview and comparative results.

Study 1		Study 2		Study 3	
Consumer e-learning Between-subjects (3 conditions) N1=16 Topic of investigation: behaviors/processes		Recommendation-based decision-making Between-subjects (2 stimuli) N2= 15 Topic of investigation: decision-making process		Video content marketing Within-subjects (3 stimuli: nano-, macro-influencer, brand) N3= 15 Topic of investigation: decision-making process	
Experimentation /Observation	In-depth discussion	Experimentation /Observation	In-depth discussion	Experimentation /Observation	In-depth discussion
Iterative searching behavior; Variety of information sources used; Similarities of searching behaviors across scenarios (e.g., scant reading, jumping video sections); Divergences of searching behaviors within scenarios (e.g., searching order)		Immediate searching for alternatives; Longer decision-making process ("real-time")		Comparison between stimuli (e.g. less professional); Only criteria based on visible stimuli are discussed	
		Chronological explanation of searching behavior; Dominant preference or do not remember the source; Inability to explain search behaviors into depth		Comparison with previous experience; Open and spontaneous discussion on criteria of importance	
		Conscious evaluation of positive/negative arguments regarding the recommendation; Artificially shortened decision-making process (shortcut towards recommended product)			

Based on these findings, we present a methodological framework for conducting qualitative experiments, including recommendations and techniques to address internal, external and interpretivist validity (see Figure 1). We also highlight the benefits and risk for interpretivist research.



**Figure 1:** Methodological framework for conducting qualitative experiments.

Based on the findings of our investigation, we observe that qualitative experiments are a recommended research approach for qualitative research aiming to explore and identify patterns, processes or behaviors. As participants are sometimes unable to express, or even wrongly describe, their decision-making process, qualitative experiments not only help to observe participants' real behaviors in controlled study settings but also compare their behaviors in different settings while controlling for alternative explanations. However, given the imposed study context, researchers should

be aware of directing behaviors, especially in within-subjects experiments, which become therefore less recommended for qualitative research given the reduction of interpretivist validity. More generally, we therefore recommend qualitative experiments only for the exploration of patterns, processes and behaviors, since other research topics commonly investigated in qualitative research require a more grounded and interpretivist approach to uncover feelings, motivations, etc. When designing qualitative experiments, our findings recommend addressing internal, external and interpretivist validity using the variety of techniques identified in the literature and empirically tested across our three studies (see Table 2).

## **6. Conclusions**

This research offers a primer view of modern qualitative experiments combining qualitative and quantitative methods to assess validity criteria in both methodological research streams. More precisely, this paper suggests an alternative technique to explore patterns using an open but controlled research environment. The findings of this research contribute methodologically to the works of previous authors who first introduced qualitative experiments in the scientific literature (Bartlett, 1953, 1995; Kleining, 1986; Robinson & Mendelson, 2012). In particular, our findings contribute by offering a methodological framework and recommendations based on three validity criteria. While previous research either focused on quantitative validity criteria (e.g., Slack & Draugalis, 2001) or qualitative validity criteria (e.g., Whittemore, Chase, & Mandle, 2001), we highlight the importance of integrating validity criteria from both research approaches to suggest a valid study protocol and thus more reliable results. In particular, the literature review led us to present and test three types of validity criteria: internal, external and interpretivist criteria. The methodological framework connects design- and sampling-related methods and techniques to address each type of validity.

Furthermore, the findings empirically identify the risks and benefits of qualitative experiments for qualitative research. In contrast with traditional methods such as in-depth interviews or focus groups, qualitative experiments offer a complementary data collection method that helps in identifying patterns or processes that participants might be unable to express. They also enable the observation of similarities and differences across different study contexts, thanks to a rigorous study protocol. This paper thereby supports previous works' recommendation through empirical evidence (e.g., Kleining & Witt, 2000; Wagoner, 2015).

The findings are thus of interest for qualitative researchers who seek to investigate patterns or behaviors in social sciences. When using qualitative experiments, we recommend preparing a well-thought study protocol to increase internal validity, and using real tasks with a diversified sample to increase external validity. Moreover, this research insists on an open study approach regarding data collection and analysis to uncover processes and behaviors in an exploratory fashion. This increasing interpretivist validity helps identify credible and authentic insights as recommended in the literature (Whittemore, Chase, & Mandle, 2001).

Finally, this research presents some limitations. Given the heterogeneous nature of qualitative research topics in social sciences, all methodological recommendations may not be applicable to specific research topics that could not be explored in this paper.



Moreover, the use of qualitative experiments requires a minimum of scientific understanding about the study topic to prepare a relevant research design. The use of qualitative experiments becomes thus difficult in nascent research streams.

## 7. References

- Andrews, I. & Oster, E. (2018). Weighting for external validity. NBER Working paper series, N°23826, 1-62.
- Bartlett, F. C. (1952). Review of thinking: An introduction to its experimental. *Quarterly Journal of Experimental Psychology*, 4(1), 87-90. <https://doi.org/10.1037/h0053080>
- Bartlett, F. C. (1995). *Remembering: A study in experimental and social psychology*. Cambridge University Press. (Original work published 1932)
- Campbell D. T., & Stanley J. C. (1963). *Experimental and quasi-experimental designs for research*. Houghton Mifflin.
- Cho, Y. -J., Fu, P. -W. & Wu, C. -C. (2017). Popular research topics in marketing journals, 1995-2014. *Journal of Interactive Marketing*, 40, 52-72. doi:10.1016/j.intmar.2017.06.003
- Decuir-Gunby, J., & Schutz, P. (2017). Chapter 6 mixed methods designs: frameworks for organizing your research methods. In J. DeCuir-Gunby, & P. Schutz (Eds.), *Developing a mixed methods proposal: A practical guide for beginning researchers* (pp. 83-106). Sage.
- Dülmer, H. (2015). The factorial survey: Design selection and its impact on Reliability and internal validity. *Sociological Methods & Research*, 45(2) 1-44. <https://doi.org/10.1177/0049124115582269>
- Hayashi, P. Jr., Abib, G., and Hoppen, N. (2019), Validity in qualitative research: A processual approach. *The Qualitative Report*, 24(1) 98-112. <https://nsuworks.nova.edu/tqr/vol24/iss1/8/>
- Kleining, G. (1986). Das qualitative experiment [The qualitative experiment]. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 38, 724-750. <https://www.ssoar.info/ssoar/handle/document/863>
- Kleining, G. (1991). The qualitative experiment. In U. Flick, E. V. Kardoff, H. Keupp, L. V. Rosenstiel, & S. Wolff (Eds.), *Handbuch qualitative Sozialforschung: Grundlagen, Konzepte, Methoden und Anwendungen* (pp. 263-266). Beltz.
- Kleining, G. & Witt H. (2000). The qualitative heuristic approach: A methodology for discovery in psychology and the social sciences. *Rediscovering the method of introspection as an example*. *Forum: Qualitative Social Research*, 1(1), 1-6. <https://www.qualitative-research.net/index.php/fqs/article/view/1123/2495>
- Kolb, D. A., Boyatzis, R. E. & Mainemelis, C. (2002). Experiential learning theory: Previous research and new directions. In R. J. Sternberg & L. F. Zhang (Eds.). *Perspectives on cognitive, learning, and thinking styles* (pp. 227-248). Lawrence Erlbaum Associates.

- Leviton, L. C. (2017). Generalizing about public health interventions: A mixed-methods approach to external validity. *Annual Review of Public Health*, 38(1), 371-391. <https://doi.org/10.1146/annurev-publhealth-031816-044509>
- Lincoln, Y. S. (1995). Emerging criteria for quality in qualitative and interpretive research. *Qualitative Inquiry*, 1(3), 275-289. <https://doi.org/10.1177/107780049500100301>
- Lynch, J. G. (1982). On the external validity of experiments in consumer research. *Journal of Consumer Research*, 9(3), 225-239. <http://www.jstor.org/stable/2488619>
- Lynch, J. G. (1983). The role of external validity in theoretical research. *Journal of Consumer Research*, 10(1), 109-111. <https://www.jstor.org/stable/2488860>
- Maxwell, J. A. (1996). *Qualitative research design: An interactive approach*. Sage.
- Ramirez-Montoya, M.-S. & Lugo-Ocando, J. (2020). Systematic review of mixed methods in the framework of educational innovation. *Media Education Research Journal*, 28(65) 9-20. <https://eric.ed.gov/?id=EJ1267414>
- Robinson, S. & Mendelson, A. L. (2012). A qualitative experiment: Research on mediated meaning construction using a hybrid approach. *Journal of Mixed Methods Research*, 6(4), 332-347. <https://doi.org/10.1177/1558689812444789>
- Sandelowski, M. (1986). The problem of rigor in qualitative research. *Advances in Nursing Science*, 8(3), 27-37. <https://doi.org/10.1097/00012272-198604000-00005>.
- Seltman, H. J. (2018). *Experimental design and analysis*, e-book <http://www.stat.cmu.edu/~hseltman/309/Book/Book.pdf>
- Slack, M.K. & Draugalis, J.R. (2001). Establishing the internal and external validity of experimental studies. *American Journal of Health-System Pharmacy*, 58(22), 2173-2181. <https://pubmed.ncbi.nlm.nih.gov/11760921/>
- Wagoner, B. (2015). Qualitative experiments in psychology: The case of Frederic Bartlett's methodology. *Forum: Qualitative Social Research*, 16(3), 1-38. <https://www.qualitative-research.net/index.php/fqs/article/view/2367/3870>
- Whittemore, R., Chase, K.C., and Mandle, C.L. (2001). Validity in qualitative research. *Qualitative Health Research*, 11(4), 522-537. <https://doi.org/10.1177/104973201129119299>