International Journal of Educational Methodology

Defining Preliminary Research for Digital Game-Based Learning Evaluation: Best Practices

Denise Pauline-Graf *
Novomatic AG, AUSTRIA

Susan Elaine Mandel
Lake Health, USA

Abstract: The surge of learners being immersed in computer game contexts for learning has instigated dialogue about the contextually appropriate collection of reliable and valid data to inform education-based decisions. The purpose of this article is to develop educational practitioners' understanding of preliminary research work, and to inform educational researchers about design and reporting of preliminary research work, in the context of reported preliminary studies on Digital Game-Based Learning Interventions (DGBLIs). First a checklist of processes for the reporting of preliminary studies is provided. Second, a summary is offered of the characteristics of each type of preliminary study including the description, objectives, and methodology. Third, an example from peer-reviewed literature is identified of each type of preliminary study relevant to DGBLIs and conducted within the past five years. Evident from the examples selected, educational researchers and practitioners are best advised to recognize the characteristics of preliminary studies — pilot work, feasibility study, pilot study, pilot trial, and field test — to better inform DGBLIs before embarking on a full-scale study, and to meet the need of educational practitioners for concrete evidence about DGBLIs.

Keywords: *Preliminary research studies, digital game-based learning interventions, pilot study, field test, education methodology.*

Introduction

Technologies, as tools to produce significant changes in learning outcomes, are emerging in schools. To attain enhanced effective learning of people, institutions require instructional strategies that include a focus on the learner, which allows them to gain knowledge by doing (Petri & Gress von Wangenheim, 2017). Digital game-based learning, which incorporates educational content into computers with the aim of engaging learners in an entertaining way (Connolly, Boyle, MacArthur, Hainey, & Boyle, 2012), is an innovative instructional method to achieve diverse educational outcomes (Cowley, Fantato, Jennett, Ruskov, & Ravaja, 2014; Petri & Gress von Wangenheim, 2016) including academic skills (Main, O'Rourke, Morris, & Dunjey, 2016) efficiency behaviors (Reeves, Cummings, Scarborough, & Yeykelis, 2015), emotional regulation (Astor, Adam, Jercic, Schaaf, & Weinhardt, 2013), and preventive health behaviors (Read & Shortell, 2011). However, educational researchers must properly connect their research to the challenges of everyday practice of Digital Game-Based Learning Interventions (DGBLIs) to enable educational practitioners to advance from a standpoint, in which the only proof they may possess for a DGBLI's effectiveness is that learners found it fun, toward a professional position in which their educational decisions are assiduously justified by the generation of valid and reliable study findings and from the accurate assessment of their practices specific to their teaching environment.

The increasingly prominent role computer and console games play in youth culture (Toomey 2017) and the surge of learners being immersed in computer game contexts for learning (Tuzun, Barab & Thomas, 2019) has instigated dialogue about the contextually appropriate collection of reliable and valid data to inform education-based decisions. Sims, Riley-Tilman and Cohen (2017) noted that a rising focus exists on data-based decision making to inform education in schools. The United States federal government is holding schools accountable for the education they provide, which requires evidence (Wiens, 2014). The need for more and better use of empirical evidence to inform education is rising not only in the United States but also internationally. Educational knowledge is global and abstract, and scientific proof is an effective, universal indicator of knowledge and learning (Wiseman, Whitty, Tobin & Tsui, 2010).

* **Corresponding author:**
Denise Pauline-Graf, Novomatic AG, Wiener Strasse 158, 2352 Gumpoldskirchen, Austria. ✉ denizh8@gmail.com

The purpose of this article is to develop educational practitioners' understanding of preliminary research work, and to inform educational researchers about design and reporting of preliminary research work, in the context of reported preliminary studies on DGBLIs. The first author conducted a preliminary randomized trial in a study on a knowledge sharing intervention program using instructional digital games (Paulic, 2017). In this study, preliminary research was presented to guide efforts in the design of a DGBLI for employees' experiential learning to share knowledge in organizations. The first author opted to conduct a randomized trial before embarking on a costly, full-scale study. Organizations interested in the instructional uses of knowledge-based intervention programs using digital gaming technology could take this preliminary evidence and compare it to alternate knowledge-based training interventions.

Experience with academic literature left the first author with doubt concerning preliminary research work terminology and apprehension toward the processes and methodology of the various preliminary studies. Experience with academic literature has left many educational practitioners with skepticism concerning the practical use and value of research in educational settings (Montgomery & Smith, 2015). Digital game-based learning interventions are an intricate endeavor with a direct call for educational researchers to present well-defined approaches for the systematic assessment of instructional games to obtain valid and reliable results (Petri & Gress von Wangenheim, 2016).

Practitioners hold, for example, a full-scale, randomized controlled trial as the gold standard research design to test a cause and effect relationship; however, despite its strength in minimizing bias and confounding factors (Rajasekar & Kumar, 2019) this study design may be unhelpful to practitioners who need to facilitate the testing of digital games in a practical and more digestible manner; the design is costly in resources (Rajasekar & Kumar, 2019). For example, a full scale study may consume inordinate amounts of administrative time to assess the long-term effect of a DGBLI and other logistical constraints including scope of the study or the selection and allocation of study participants to possess the same characteristics as the reference population.

If practitioners conduct and report preliminary work well, lessons learned from it are as useful in the evaluation of DGBLIs as full-scale studies. Preliminary research work assists practitioners to draw on the unique interplay of people, the game, and the learning outcome association prior to a full scale study. Even in contexts where there is use of a standardized protocol, this interplay naturally creates complexity. The characteristics of human interaction and the nature of games are intrinsically variable. Also, researchers can further develop their studies from initial findings from preliminary research work. Subsequently, preliminary research work offers value to the assessment of interventions by enhancing the skills of educational researchers and assisting in the refinement of interventions before conducting a lengthy and expensive full-scale study (Doody & Doody, 2015).

A successful investigation larger than a preliminary study requires appropriate methodology. Key components of a study's methodology include the study's (a) method (quantitative, qualitative, or mixed); (b) research design; (c) setting; (d) population of interest; (e) sampling plan, (f) recruitment strategy and informed consent process; (g) pre-activities, (h) randomization, (i) confidentiality, (j) instrumentation, (k) data collection, (l) data analysis, and (m) the post-plan including follow-up measures. In comparison to full-scale studies, preliminary studies are less costly studies that encompass a small sample to investigate parts of an intervention system instead of the system in its entirety. Preliminary studies do not involve complex statistical analyses and the results from such studies are difficult to generalize because the sampling errors are very large. Preliminary studies differ from full-scale studies in terms of sample size, study cost, study purpose, data analysis, and generalizability of results.

However, the misrepresentation and misuse of some studies in preliminary research work are common, particularly in health research and technology (Arnold et al., 2009; Doody & Doody, 2015), and are likely to occur in other fields of evidence-based practice such as education. The mislabeling of some preliminary research work, for example trials as *pilot,* may arise to disguise the use of small sample sizes or to evade conducting studies in more than one location when the intent is to run these studies with a treatment comparison as the main objective (Whitehead, Sully, & Campbell, 2014). Common errors in pilot data reporting include poor research methodology and an inappropriate focus on intervention effectiveness (Arain, Campbell, Cooper & Lancaster, 2010).

A predominant concern for educational researchers, who seek to establish whether people respond to an intervention, is the reporting of the "positive, negative, small, large, or negligible intervention's effects to guide intervention activities" (Sims et al., 2017, p. 7). Nevertheless, limited information exists to guide educational researchers on understanding the characteristics of the various studies in preliminary research work, including their descriptions, objectives, methodology, and analysis to evaluate a complex intervention for the optimal achievement of a study's purpose, perpetuating misunderstandings with educational practitioners who require concrete evidence in touch with the day-to-day realities of their educational setting. Preliminary studies, useful to inform DGBLIs in fields of evidence-based practice, include *pilot work*, *feasibility studies*, *pilot studies*, and *pilot trials* (Arnold et al., 2009). In addition to these four preliminary studies are *field tests*.

Diverse interpretations of these terms exist in the assessment of DGBLIs, often with educational researchers adopting their own definitions to adapt to their specific projects and methods. Frequently occurring examples are designers of interview schedules and evaluation questionnaires referring to 'piloting' with a small sample of participants similar to

the target population for their research (Jasem & Delport, 2019; Stouck, 2019). However, a requisite for each preliminary study is a well-defined description, a general and a specific objective, and appropriate methodology.

An alternative approach to preliminary studies, to explore and address a research question or test a DGBLI, is proposed. Unlike pilot work and feasibility studies that center on the studies' components, pilot studies, pilot trials, and field tests are defined as preliminary studies that include a focus on intervention effectiveness in addition to how the studies' components work together. Explicit criteria for the success of a carefully chosen preliminary study are a prerequisite and must be based on the preliminary study's specific objective. For example, criteria for the success of pilot work, with the objective to uncover background information on a DGBLI, will differ from criteria for the success of a pilot trial with the objective of assessing the intervention's effectiveness as well as the feasibility and acceptability of the intervention's outcome measures before conducting a full-scale study.

A study method and design must be best suited to generate qualitative or quantitative data that achieves the study's objectives. The formal evaluation of evidence to obtain conclusive results is unfitting in preliminary research work on DGBLIs; rather, focus must be on the implementation and evaluation of procedures and the analysis of preliminary findings with a small sample. The naturalistic paradigm that underpins qualitative research methods can involve researchers' interaction with the study's participants. Researchers can observe the participants, with or without joining a group, to collect highly context specific observations to gain deep insight into a research problem (Sullivan & Sargeant, 2011). The positivist paradigm that underpins quantitative research methods requires educational researchers to remain independent from the study's participants and to test a hypothesis through direct observations under controlled contexts (Sullivan & Sargeant, 2011). Formal hypothesis testing is uncharacteristic of studies in preliminary research work; instead, they may involve descriptive statistics as done in pilot studies and pilot trials (Lee, Whitehead, Jacques, & Julious, 2014). Post-hoc power calculations are possible but unfavorable in pilots; rather, researchers must infer the treatment effect's size and direction using estimations and confidence intervals. Further, a disclaimer of inadequate study power must accompany any interpreted ρ-values (Lee et al., 2014).

Mixed methods involve both qualitative and quantitative elements in a study (Guest, 2012). The pragmatic paradigm that underpins mixed methods research includes the notion that social science research does not involve a commitment to one study method. Focus is on an appropriate way to address the research problem in the natural world (Duemer & Zebedi, 2009). On a methodological level, the pragmatic approach has limitations as an output of pending philosophical problems that call for extensive and time-consuming data collection (Duemer & Zebedi, 2009) and hence inappropriate for preliminary research work on DGBLIs.

Common to digital game technology in the initial phase of fast development is the need to complete preliminary studies in a reasonable amount of time to establish whether or not to progress toward the technology's broader implementation. Hence, the concrete research objective rather than the philosophical position determines the methodology of the preliminary study. The use of mixed methods must be questioned as a genuine contribution to the preliminary study's objective, particularly when preliminary research work involves a phased approach to implementation, to optimally answer a research question. Moreover, adequate resources (human and non-human) must be available to integrate two distinct methods, one thematic (qualitative) and the other statistical (quantitative), into a single inquiry better suited for a comprehensive full-scale study. Therefore, mixed methods are plausible to merge the results of two complementary research paradigms for a full-scale study rather than a preliminary study.

In the methods section of this article, chief methodological features of the varying preliminary studies in educational digital gaming were synthesized and their value highlighted to provide practical guidelines for the design of possible future studies to inform DGBLIs.

Methodology

First, a checklist of processes for educational researchers to consider when reporting preliminary studies was provided. Second, a summary was offered of the characteristics of each type of preliminary study including the description, objectives and methodology. Third, from peer-reviewed literature an example was identified of each type of preliminary study relevant to DGBLIs and conducted within the past five years.

Reporting of a Preliminary Study: Checklist of Processes

A point of interest to educational researchers is the assessment of DGBLIs in the context of reported preliminary research work as an aid to understanding the differences among the types of preliminary research work and subsequently the evidence-based selection of digital games, ultimately influencing educational policy and instructional practice. Table 1 is a checklist of processes for the reporting of a preliminary study. Preliminary research work involves small-scale testing of an intervention and is crucial to educational researchers who seek to learn about educational outcomes as well as applicability and transferability of the intervention to other settings, populations, or time frames. Further, researchers commonly conduct preliminary studies to refine interventions (Smith, Morrow & Ross, 2015). The processes vary depending on the selection of the preliminary study type.

Table 1. Checklist to Consider when Reporting Preliminary Studies

Process Check	Checklist Item	Steps	
Background and objectives		1	Develop the research area and rationale for the research problem. Develop specific objectives, research questions, or hypotheses
Methodology	Study method	2	Select the method (qualitative or quantitative)
	Study design	3	Address the research problem by selecting the appropriate research design
	Setting	4	Obtain permissions for location access to collect data
	Population of interest	5	Determine the target population under study
	Sampling plan	6a	Define eligibility criteria for participants and determine sample size
	Recruitment strategy	6b	Obtain permissions to recruit individuals and use premises for the study
			Estimate the pool of available individuals who meet the inclusion criteria to inform the design of a full-scale study
			Calculate the recruitment rate, which serves as an indicator to either continue or abandon a full-scale study
			Recruitment rate can be calculated as the proportion of individuals contacted, who gave study participation consent and invited to complete the questionnaire
	Informed consent	6c	Detail the manner in which individuals provide informed consent to take part in the research. The informed consent form must include the study's purpose and nature, where the study takes place, and the manner in which data are obtained and recorded without adverse consequences to those who decide to participate in the study
			Inform individuals about plans for maintaining anonymity and confidentiality of information collected from the study's participants to protect privacy, and how to voluntarily withdraw from the study
	Pre-study activities	7a	Outline the intervention, including who will administer the intervention to participants or groups, and define when, where, and how the intervention will be conducted to allow for study transferability or replication
		7b	Determine how confounding variables are controlled
		7c	If a pilot study of the intervention is conducted, determine if the pilot will be internal or external to the full-scale study
	Randomization Sequence generation	8a	If randomization is applicable, determine the random sequence generation (such as using a software program or random-number's table)
		8b	Establish the randomization as simple, block, or stratified
	Allocation concealment technique	8c	Determine the technique to conceal participant allocation until recruiters have assigned participants to the interventions (such as contacting a central methods center by computer or phone)
	Implementation	8d	Determine who manages participant enrolment, group randomization, participant allocation to groups, and participant exposure to the intervention and conducts a follow-up analysis

Table 1. Continued

Process Check	Checklist Item		Steps
	Blinding	8e	If the experiment is blinded, determine who will be blinded (such as trainers, participants, those evaluating outcomes) and how blinding is achieved after participants are assigned to the interventions
	Confidentiality	9	Anonymize, encrypt, and store data so that no person can trace identifiers back to the study's participants
	Instrumentation	10	Assess and develop primary and or secondary data collection tools for data collection
	Data collection plan	11a	If the study is qualitative, use research methods that build on non-numerical data (such as individual interviews, focus groups, naturalistic observations)
		11b	If the study is quantitative, use research methods that adopt the standardization of numeric data collection to allow for statistical comparison (e.g., surveys, face-to-face interviews, systematic observations, and online polls)
	Data analysis plan	12a	Use thematic analysis to understand the phenomenon under study or statistical analysis to determine within-group or between-group statistical differences for primary and secondary outcome measures
		12b	Use subgroup analyses and adjusted analyses for additional evaluations
	Post study plan	13	Make essential changes to the study's design after the study's commencement and decide how participants learn about the research

Evident from Table 1, researchers need to determine the appropriateness of the preliminary study's processes before undertaking a full-scale study. The processes vary depending on the type of preliminary research study the researcher selects. For instance, feasibility study processes will differ to processes of a pilot trial or field test of a DGBLI.

Characteristics of Preliminary Studies

In an effort to promote methodological rigor and the publication of empirically valid papers on preliminary research work for the assessment of DGBLIs, examples from peer-reviewed literature on the characteristics of preliminary studies are described as follows:

Pilot work—"Background research that informs a future study" (Arnold et al., 2009, p. 69). The goal of pilot work is to gain background information to understand the rationale behind some aspects of the full-scale study and precedes feasibility studies, pilot studies, or pilot trials. Pilot work can involve the collection and interpretation of data from participants using averages to obtain input on the ease of use and usefulness of an educational digital game as effectively demonstrated by Wood, Mentzelopolous, and Protopsaltis (2015) pilot work of a digital game developed to assist undergraduates with the learning of programming. The next phase to the pilot work can include the assessment of the effectiveness of the learning outcome using pre- and post-test questionnaires (Wood et al., 2015).

Feasibility studies—Preliminary studies that simulate aspects of pilot studies but do not exhibit the scope of pilot studies that emulate an intended full-scale study (Whitehead et al., 2014). The goal of feasibility studies is to assess a study's viability prior to conducting a full-scale study; feasibility studies may or may not involve the intention to the next phase of the study — the roll-out of a full-scale study (Whitehead et al., 2014). Feasibility studies may include meta-analysis or a systematic review (Whitehead et al., 2014).

Although feasibility studies might involve educational researchers assessing the acceptability and feasibility of an intervention's outcome measures, focus must be placed only on the evaluation of feasibility measures. Feasibility measures include the aspect of practicality including delivery, time, and costs (Golding, Kneebone, & Fife-Schaw, 2016). Parameters of delivery may include number of people eligible for the research (i.e., inclusion and exclusion criteria on recruitment efficiency); the informed consent process for human subjects (e.g., which risks must be disclosed and additional protection measures for vulnerable populations); and the recruitment process (standard deviation of the outcome measure, essential in some cases for sample size estimation, willingness of educational practitioners to recruit participants and willingness of participants to be recruited). Parameters of time may involve data availability or limitations of a certain database and the time necessary to collect and analyze the data. Parameters of costs may include the costs incurred to roll out the study. A feasibility study on the attitudes toward the practice of game-based learning

(Vate-U-Lan, 2015) must not lack focus on the estimation of parameters of delivery, time, and costs for the design of a full-scale study.

Pilot studies—Mini-versions of the intended main study (Doody & Doody, 2015) and therefore involve the recruitment of a small group of participants similar to those enrolled for an intended full-scale study. Pilot studies are kinds of feasibility studies (Whitehead et al., 2014) and therefore comprise features of feasibility studies, which include aspects of practicality including delivery, time, and costs. However, feasibility studies are "pieces of research" conducted before a larger study "to answer the question 'Can this study be done?'" (Eldridge et al., 2016, p. 9) and involve parameter estimations essential to design a full-scale study.

Pilot studies and feasibility studies both include a focus on whether something can be done, and if so, how to proceed with it; however, pilot studies have a further dimension (Eldridge et al., 2016). They include a specific research problem or hypothesis and methodology (Arnold et al., 2009). Characteristics that distinguish pilot studies from pilot work and feasibility studies are that pilot studies (a) possess a more rigid study methodology, (b) are smaller versions of full-scale studies, (c) include an emphasis on pilot study processes and whether the full-scale study's components work in unison, as well as (d) underscore the need for future definitive work (Whitehead et al., 2014). A stricter study methodology is characteristic of most pilot studies (e.g., sample size justification) compared to pilot work or feasibility studies. The more rigorous methodology stems from the fact that pilot studies are more likely to resemble the design of a full-scale study.

Pilot studies can involve the assessment of an intervention's outcome measures prior to a full-scale study and include testing the feasibility (delivery, time, and costs) and acceptability (participant recruitment rate and drop-out rate) of the outcome measures. Accordingly, "pilot studies are feasibility studies" (Whitehead et al., 2014, p. 132), but feasibility studies are not necessarily pilot studies. Similar to pilot work and feasibility studies, pilot studies can involve qualitative or quantitative data collection methods (Doody & Doody, 2015). Interviews or focus groups; written narratives; written responses to open-ended questions; nonparticipant or participant observations, and review of documents are examples of qualitative methods whereas quantitative methods involve direct observations that are controlled or assumed stable (Sullivan & Sargeant, 2011).

An emphasis on study processes is inherent in pilot studies to address specific research design features and alleviate problems in advance of a full-scale study, such as a qualitative phenomenological study (Kim, 2011) or a quantitative randomized study (Eldridge et al., 2016). Pritchard and Whiting (2012) highlighted the importance of extending the scope of pilot studies in qualitative research beyond data collection to evaluate entry and orientation to the research site as well as assess transcription and analysis procedures. Pilot study processes involve considering if the pilot study is internal or external to a full-scale study (Whitehead et al., 2014).

An internal pilot study is part of the full-scale study, popular to address uncertainties in sample size calculations (Schneider, Schmidli, & Friede, 2013). Accordingly, one can define an external pilot as a pilot study not intended to be part of the full-scale study as illustrated in Blekken et al.'s (2018) pilot study on the feasibility, acceptability, and adherence of an educational intervention program for care staff. Researchers can determine parameters such as participant recruitment rate within the full-scale study. The recruitment rate can have a specified time point, which serves as an indicator for either the continuance or abandonment of the full-scale study depending on the number of participants recruited (Whitehead et al., 2014).

Researchers decide when to proceed or abandon the full-scale study by specifying what size difference in the study's outcome will justify further extension of the pilot to include this outcome (Whitehead et al., 2014). In an internal pilot, researchers squander participants if participants abandon the pilot before it contributes to the full-scale study's end-point. Further, an internal pilot roll-out poses a challenge regarding resources such as funding, if the full-scale study's outcome is reliant on recruitment rate and appears unrelated to the participants' benefit (Whitehead et al., 2014). In an external pilot, participants are also squandered if the pilot does not contribute to the full-scale study's end-point, which threatens research validity when participants are difficult to source (Whitehead et al., 2014). Funding agencies are more likely to fund an external pilot study and include contingency funds for the full-scale study than to commit to funding an internal pilot (where the pilot is part of the full-scale study) with withdrawal of contingent funding if progress of the whole study is unsatisfactory (Whitehead et al., 2014).

To better meet the goal of pilot studies, reporters of pilot studies must consider presenting both primary and secondary outcome measures of the intervention in preparation of a full-scale study. For example, in an exploratory design of a multi-sensory computer game prototype, Serret et al. (2017) included evidence of the game's effectiveness as a tool to teach emotions to young patients with Autism Spectrum Conditions. However, the study lacked consideration toward better defining the intervention's acceptability and feasibility measures, in preparation of a full-scale, definitive qualitative or quantitative study. Clear intention in pilot studies for future definitive work is vital; otherwise, pilot studies can be underpowered with reduced scientific use; thus, mitigating unethical research practice (Doody & Doody 2015).

Pilot trials—Stand-alone pilot studies with a randomization procedure (Arnold et al., 2009) and are designed to evaluate primary as well as secondary outcome measures of an intervention in preparation for a full-scale study. Reports of pilot trials, such as Bingham, Lahiri and Ashikga's (2012) within-subjects, randomized cross-over design of digital spirometer (breath biofeedback) games, included evidence of both primary and some secondary outcome measures. Primary outcome measures included evidence of the games' effectiveness to teach adolescents with cystic fibrosis to control their inspiratory and expiratory breathing and some secondary measures involved acceptability of the intervention, including the percentage of days, patients used the software compared to the number of days the software was available. However, to better meet the goal of pilot trials, secondary measures must also include the feasibility of the intervention's outcome measures.

Field tests—Studies similar to, but more extensive than pilot studies and pilot trials, and occur at the final stage of an intervention's development in preparation of a full-scale study. Although laboratory studies include benefits such as easier randomization of participants between the intervention and control group and better stimulus control than field tests, field tests occur in natural environments and have the upside of not being isolated from real life experiences of participants (Franzen & Pointner, 2013). Richard and Belanger (2018) noted that field access involves processes to gain and maintain entry to a setting or to establish working relations with people.

Results from field tests are more likely to be generalized to real world behavior than results from laboratory settings (Franzen & Pointner, 2013). Researchers can draw upon the knowledge and skills of agents, such as expert respondents, in a collaborative, resource-intensive operation (Alber, Watson, Barnett, Mercado, & Bernhardt, 2015). Unlike pilot studies and pilot trials, which are small-scale tryouts of a full-scale intervention under development, field testing involves a more diverse population (Randel & Clark, 2013) such as in Katajapuu et al.'s (2016) quantitative field test of a digital game-based brain trainer exercise game to measure brain activity on elderly populations from a Finish university and a Japanese well-being center. Although the researchers presented results of the game's effectiveness and acceptability of the game including usability and enjoyment in game play, it lacked focus on the intervention's feasibility outcome measures.

Table 2 includes a presentation of the characteristics of each preliminary study to facilitate their identification in DGBLI studies. Disagreement about definitions may not influence the quality of preliminary research work; however, problems may arise when educational researchers attempt to develop guidance in the form of standardized procedures for preliminary research conduct. Clarification of what the procedures apply to and subsequently what steps they involve is vital for the conduct of preliminary research work for DGBLIs.

Table 2. Characteristics of Preliminary Studies

Characteristic	Pilot work	Feasibility Study	Pilot study	Pilot trial	Field test
Description	Background information to inform a study of interest	Preliminary study that simulates aspects of a pilot	Kind of feasibility study that emulates an intended full-scale study The pilot study can be internal or external to a full-scale study	Randomized, stand-alone pilot study	Similar to pilot studies and trials but occurs on a larger scale and under conditions of actual use or operation
Objectives	Gain information about a study of interest in preparation of a full-scale study. Excludes a focus on problem alleviation	Assess study viability before conducting a larger study. Need not include the intention to conduct further work	Evaluate a study's outcome measures as well as feasibility and acceptability of the study's outcome measures in preparation of a full-scale study	Evaluate a study's outcome measures as well as feasibility and acceptability of the study's outcome measures in preparation of a full-scale study	Evaluate a study's outcome measures as well as feasibility and acceptability of the study's outcome measures under conditions of actual use or operation, in preparation of a full-scale study

Table 3. Continued

Characteristic	Pilot work	Feasibility Study	Pilot study	Pilot trial	Field test
Methodology	Minimal methodology rigor	A more rigorous methodology than pilot work	A more rigorous methodology than feasibility studies but less than a full-scale study	A more rigorous methodology than feasibility studies but less than a full-scale study	A more rigorous methodology than feasibility studies but less than a full-scale study
	Qualitative or quantitative	Qualitative or quantitative	Qualitative or quantitative	Quantitative	Qualitative or quantitative
Example	Identifying features and conditions of a study of interest	Testing the viability of a study of interest	Testing aspects of a quantitative or qualitative study in preparation of a full-scale study assessment	Testing aspects of a quantitative study in preparation of a full-scale study assessment	Testing aspects of a qualitative or quantitative study under conditions of actual use or conditions in preparation of a full-scale study assessment

Evident from Table 2, a succinct description, suitable objectives, and an appropriate methodology are suggested for each preliminary study for the optimal design and reporting of a preliminary study relevant to DGBLIs in different learning areas.

Findings

Emerging technology is a key aspect of learning for organizations and will continue to surge in the foreseeable future. Digital game-based learning interventions are advantageous because they expose participants to environments and experiences not usually possible in the real-world (Procci, Lakhmani, Hussain & Bowers, 2014) at variable costs lower than traditional instructional mediums (Strader, 2011). We replaced old methodological assumptions with new ones that require thoughtful scholarly discussion to reach consensus on the characteristics of preliminary research for DGBLIs.

Evident from the examples selected on preliminary research work in the context of DGBLIs, educational researchers and practitioners are best advised to recognize the characteristics of preliminary studies — *pilot work*, *feasibility study*, *pilot study, pilot trial,* and *field test* — to better inform DGBLIs before embarking on a full-scale study, and to meet the need of educational practiters for concrete evidence about DGBLIs. The initiation of best practices to produce superior preliminary study outcomes are reliant on the expertise and experience of educational researchers who implement a standardized protocol for the assessment of DGBLIs that can be conducted flexibly across diverse sectors and clearly understood by educational practitioners.

Discussion and Conclusion

Time and effort is necessary to understand the complexity of DGBLIs. This article is a start to the formal clarification of preliminary studies in preliminary research work to develop critical thinking about preliminary research work, a crucial cognitive skill for educational practitioners who are called to critically assess, analyze and reasonably deliberate on the intervention's potential to prepare learners for self-direction. Educational practitioners can form partnerships with educational researchers as a means to promote more research specific to their learning environments. Educational practitioners can draw on the expertise of educational researchers to design preliminary studies on DGBLIs that help them answer their personal queries about teaching and learning with games. As educational practitioners start to collaborate and affect the purpose for which researchers gather data, they can exert an erudite influence on the direction of research on DGBLIs. Educational practitioners' relationships with research can equip them to better assess claims concerning the latest trends in teaching methodologies using digital games, and can support them in justifying and defending their instructional decisions to policy-makers, administrators, colleagues, and parents.

Accordingly, educational researchers who partner with practitioners can gain access to opportunities to examine DGBLIs in varying environments. Drawing on the expertise of practitioners, educational researchers must design meaningful preliminary studies that are more likely to address issues of concern to practitioners. Educational researchers must be prepared to expect both favorable and unfavorable results from DGBLIs by thinking through in

advance the inputs and outcomes of such a complex system and search for ways to adequately replicate the results. Preliminary research work helps to address issues in each phase of the intervention's development and implementation before investing resources in a full-scale study. Therefore, a checklist of the processes for the reporting of preliminary studies and a clear distinction among the characteristics of preliminary research studies, including description, objectives and methodology must be universally observed.

This article includes six points of discussion about the need for educational researchers to reach consensus on the characteristics of preliminary studies. First, diverse views exist about the definition of preliminary studies for DGBLIs, which can result in ambiguous interpretations of inputs and outcomes. From an etymological point of view, definitions require revision and amendment as knowledge evolves and practice progresses. For instance, the term pilot first emerged as meaning something is a prototype of something else (Bertoni, Bordegoni, Johansson, & Larson, 2008)**.** By mid-20th century the term was associated with research design, as 'pilot survey' (Eldridge et al., 2016). Until educational researchers gather information that facilitates the development of predictive text, unfamiliarity with DGBLIs will be associated with ambiguity. Educational researchers must seek expert counsel to clarify definitions related to preliminary research work to better inform DGBLIs in diverse educational contexts. For example, the phrase *pilot testing* used in Wood et al.'s (2015) pilot work must rather be used in preliminary studies that involve pilot testing a hypothesis such as in pilot trials. Mutually exclusive definitions for the various preliminary studies are advised to determine whether or not the preliminary study is planned in preparation of a full-scale study of the DGBLI.

Second, inconsistent term usage is common in preliminary research work (Eldridge et al., 2016). Educational researchers must not use terms for preliminary research studies interchangeably such as the use of the term feasibility study in a pilot trial as discovered in Bingham et al.'s (2012) pilot trial of spirometer games, or the use of the term *pilot study* in a pilot trial as found in Boendermaker et al.'s (2018) research on motivational game elements for cognitive training. Such studies lead to the omission of important study parameters to better inform DGBLIs.

Third, educational researchers must not use terms imprecisely. The use of the term *pilot study* is more appropriate than *feasibility study* to explore the attitudes of teachers, parents, and university students toward digital game-based education (Vate-U-Lan, 2015). Feasibility studies include a focus on the parameters of delivery, time, and costs of the study whereas pilot studies are miniature versions of the full-scale study and must include sufficient assurance to enable a full-scale study.

Fourth, the role each preliminary study plays is often ambiguous. Some studies include an emphasis on goals of cause and effect, despite advice that pilot work and feasibility studies not center on the effect of the intervention. Wood et al.'s (2015) pilot work is an effective example of the analysis of data using means. However, Vate-U-Lan's (2015) feasibility study that involves qualitative and quantitative survey questions to investigate attitudes toward the practice of game-based learning must rather include the question about "whether a study can be done" (Bugge et al., 2013, p. 353). Feasibility studies need not be considered a subset of pilot studies, pilot trials, or field tests and should only include feasibility of the outcome measures. Pilot studies, pilot trials and field tests play a pivotal role in the evaluation of DGBLIs and in the planning of full-scale studies. Educational researchers must separate measures of these studies into clear primary and secondary outcome measures. Secondary outcome measures including feasibility and acceptability of the outcome measures must be two different concepts. Feasibility objectives of the outcome measures include aspects of study practicality including delivery, time, and costs. Acceptability objectives of the outcome measures include aspects of study satisfaction including recruitment rate and drop-out rate before conducting a full-scale study.

Fifth, preliminary studies must use the appropriate methodology. Mixed methods are more appropriate for a full-scale study instead of a preliminary study that does not warrant extensive data collection, such as Janssen and colleagues' (2015) pilot study in which they used "mixed methods to pilot a team-based digital game for boosting medical student engagement with anatomy and histology content" (p. 1). Pilot trials must also differ from other preliminary studies in that they are stand-alone pilot studies with a randomization procedure.

Sixth, incumbent upon educational researchers interested in the exploration of DGBLIs is consensus on the characteristics of preliminary studies to inform future research of DGBLIs before embarking on a lengthy, expensive full-scale study. The labeling of studies of preliminary research work must align with the study's description, objectives, and methodology preferably in both the title and abstract of the study. This suggestion applies to all studies in preliminary research work that contribute to the evaluation of a DGBLI, including qualitative assessments. The use of terms in the title and abstract of a preliminary study of DGBLIs, relevant to the setting and population, will be beneficial to both educational researchers and practitioners conducting future electronic searches. Studies must have clear descriptions of the objectives and methodology and include reports of whether they are in preparation of a full-scale study. Experienced educational practitioners can assist educational researchers to bolster their analysis of data and assure the validity of their interpretations.

Suggestions

We examined existing methodologies in preliminary research work involving DGBLI, what Barney (2018) noted as normal science and rejected the logic of some prior work; revolutionary science. In practice, issues can arise that

prohibit the implementation of the recommended characteristics of the different preliminary studies including description, objectives, and methodology. Nevertheless, in addition to this study's references, researchers are urged to seek resources to advance uniformity when defining and reporting preliminary studies for the publication of empirically valid papers. An avenue for further research would involve validating these recommendations by comparing differing preliminary study methodologies in published and unpublished research on DGBLIs (to account for publication bias) to the best practices proposed in this article and implementing these recommendations in different settings (e.g., educational institutions, healthcare facilities, and workplace). Correct implementation of these studies does not only inform full-scale studies but also improves study transferability in varying contexts.

Further, preliminary research work assists educational researchers and practitioners to build bridges among the preliminary studies' outcomes so that they can make connections between empirical studies on the role of DGBLIs and methodology as well as identify educational factors that influence persons and populations. Collaboration between educational researchers and practitioners can assist researchers to craft more compelling arguments that ultimately affect instructional training in more significant ways.

The roll-out and reporting of well-designed preliminary studies will enhance methodological rigor of empirical research on DGBLIs, thereby preempting implementation challenges common to these interventions in distinct contexts. Concerns related to DGBLIs and their implementation include research method and design appropriateness; research questionnaire adequacy; adequacy of the participant recruitment strategy; effectiveness of participant retention practices; realistic execution of the research protocol; the strength of key variable relationships; control of the confounding variables; adequacy and representativeness of the sample obtained. Other concerns are determination of resources including the number of training staff, type of training materials, as well as the time necessary to roll out the intervention, administer questionnaires, receive completed questionnaires from participants; as well as the efficient and effective analysis of data. Additional concerns involve the DGBLI's effectiveness, such as the need to visit dimensions that support game design to avoid increasing costs of an inordinately resourced full-scale study. Costs can involve hardware costs, labor costs of trainers, and opportunity costs of trainees away from their usual routine to engage in a digital game. Preliminary evidence can contribute to the understanding of user profiling (gaming experience and computer skills of a user); the context within which behavioral competencies take place (the place of learning including atmosphere); selection of learning competency pedagogies (learning framework method and game mechanics), and mode of representation (game concept, game engine, mode of deployment, and level of interactivity).

Last, educational researchers responsible for the welfare of study volunteers, are scientifically and ethically obligated to attempt to publish preliminary research work in advance of a full-scale study (Conn, Algase, Rawl, Zerwic, & Wyman, 2010; Wolfe, 2013). Kerr (2013) predicted a switch from traditional textbooks toward a more adaptive and interactive learning context with the utilization of data and analytics to integrate gamified aspects and store users' details. Educational researchers must therefore optimize their recommendations for future preliminary research work to bridge the gap that currently exists between educational research and practice.

This article does not only offer clear explanations of types of preliminary research studies that are appropriate for initial testing of DGBLIs but also of other computer-based interventions in educational contexts. Knowledge gained from the preliminary research work on both the advantages and disadvantages of DGBLIs will aid educational researchers and practitioners to test ideas and generate specialized knowledge not only on diverse aspects of digital games but also on other forms of computer-based interventions in contexts that benefit teaching and learning.

Acknowledgements

References

Alber, J., Watson, A., Barnett, T., Mercado, R., & Bernhardt, J. (2015). Development of a coding instrument to assess the quality and content of anti-tobacco video games. *Cyberpsychology, Behavior, and Social Networking*, *18*(7), 417-425. https://doi.org/10.1089/cyber.2015.0051

Arain, M., Campbell, M., Cooper, C., & Lancaster, G. (2010). What is a pilot or feasibility study? A review of current practice and editorial policy. *BMC Medical Research Methodology*, *10*(1), 67-74. https://doi.org/10.1186/1471-2288-10-67

Arnold, D., Burns, K., Adhikari, N., Kho, M., Meade, M., & Cooke, D. (2009). The design and interpretation of pilot trials in clinical research in critical care. *Critical Care Medicine*, *37*(1 Suppl), 69-74. https://doi.org/10.1097/CCM.0b013e3181920e33

Astor, P., Adam, M., Jercic, P., Schaaf, K., & Weinhardt, C. (2013). Integrating biosignals into information systems: A NeuroIS tool for improving emotion regulation. *Journal of Management Information Systems*, *30*(3), 247-277. https://doi.org/10.2753/MIS0742-1222300309

Barney, J. (2018). Editor's comments: Theory contributions and the AMR review process. *Academy of Management Review*, *43*(1), 1-4. https://doi.org/10.5465/amr.2017.0540

Bertoni, M., Bordegoni, M., Johansson, C., & Larsson, T. (2008). Pilot specifications definition guidelines for the implementation of a KEE solution in the aeronautical domain. In: Fred J. A. M. van Houten (Ed.), *CIRP Design Conference* (pp.1-7). Enschede, The Netherlands: University of Twente.

Bingham, P., Lahiri, T., & Ashikaga, T. (2012). Pilot trial of spirometer games for airway clearance practice in cystic fibrosis. *Respiratory Care*, *57*(8), 1278-1284. https://doi.org/10.4187/respcare.01263

Blekken, L., Nakrem, S., Gjeilo, K., Norton, C., Moekved, S., & Vinsnes, G. (2015). Feasibility, acceptability, and adherence of two educational programs for care staff concerning nursing home patients' fecal incontinence: a pilot study preceding a cluster-randomized controlled trial. *Implementation Science*, *10*(1), 1-13. https://doi.org/10.1186/s13012-015-0263-8

Boendermaker, W., Gladwin, T., Peeters, M., Prins, P., & Wiers, R. (2018). Training working memory in adolescents using serious game elements: Pilot randomized controlled trial. *JMIR Serious Games*, *6*(2), e10. https://doi.org/10.2196/games.8364

Bugge, C., Williams, B., Hagen, S., Logan, J., Glazener, C, Pringel, S., & Sinclair, L. (2013). A process for decision-making after pilot and feasibility trials (ADePT): Development following a feasibility study of a complex intervention for pelvic organ prolapse. *Trials*, *14*, 1-29. https://doi.org/10.1186/1745-6215-14-353

Conn, V., Algase, D., Rawl, S., Zerwic, J., & Wyman, J. (2010). Publishing pilot intervention work. *Western Journal of Nursing Research*, *32*(8), 994-1010. https://doi.org/10.1177/0193945910367229

Connolly, T., Boyle, E., MacArthur, E., Hainey, T., & Boyle, J. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education*, *59*(2), 661-686. https://doi.org/10.1016/j.compedu.2012.03.004

Cowley, B., Fantato, M., Jennett, C., Ruskov, M., & Ravaja, N. (2014). Learning when serious: Psychophysiological Evaluation of a Technology-Enhanced Learning Game. *Educational Technology & Society*, *17*(1), 3-16.

Doody, O., & Doody, C. (2015). Conducting a pilot study: case study of a novice researcher. *British Journal of Nursing*, *24*(21), 1074-1078. https://doi.org/10.12968/bjon.2015.24.21.1074

Duemer, L., & Zebedi, A. (2009). The pragmatic paradigm: An epistemological framework for mixed methods research. *Journal of Philosophy and History of Education*, *59*, 163-168.

Eldridge, S., Lancaster, G., Campbell, M., Thabane, L, Hopewell, S., Coleman, & Bond, C. (2016). Defining feasibility and pilot studies in preparation for randomised controlled trials: Development of a conceptual framework. *PLOS ONE*, *11*(3), 1-22. https://doi.org/10.1371/journal.pone.0150205

Franzen, A., & Pointner, S. (2013). The external validity of giving in the dictator game: A field experiment using the misdirected letter technique. *Experimental Economics*, *16*(2), 155-169. https://doi.org/10.1007/s10683-012-9337-5

Golding, K., Kneebone, I., & Fife-Schaw, C. (2016). Self-help relaxation for post-stroke anxiety: a randomised, controlled pilot study. *Clinical Rehabilitation*, *30*(2), 174-180. https://doi.org/10.1177/0269215515575746

Guest, G. (2012). Describing mixed methods research: An alternative to typologies. *Journal of Mixed Methods Research*, *7*(2), 141-151. https://doi.org/10.11771558689812461179

Janssen, A., Shaw, T., Goodyear, P., Kerfoot, B. P., & Bryce, D. (2015). A little healthy competition: using mixed methods to pilot a team-based digital game for boosting medical student engagement with anatomy and histology content. *BMC Medical Education*, *15*(1), 1-10. https://doi.org/10.1186/s12909-015-0455-6

Jasem, Z., & Delport, S. (2019). Mother's perspectives on the play of their children with Attention Deficit Hyperactivity Disorder. *Occupational Therapy International*, *2019*, 1-10. https://doi.org/10.1155/2019/6950605

Katajapuu, N., Granholm, P., Hiramatsu, M., Ishihara, E., Hirayama, J., Pitkakangas, P. ... Luimula, M. (2016). Brain trainer exercise game. Field tests in Finland and Japan. *International Journal of Chemistry and Chemical Engineering Systems*, *1*, 39-45.

Kerr, D. (2013). Identifying key features of students' performance in educational video games and simulation through cluster analysis. *Journal of Educational Data Mining*, *4*(1), 144-182.

Kim, Y. (2011). The pilot study in qualitative inquiry: identifying issues and learning lessons for culturally competent research. *Qualitative Social Work*, *10*(2), 190-206. https://doi.org/10.1177/1473325010362001

Lee, E., Whitehead, A., Jacques, R., & Julious, S. (2014). The statistical interpretation of pilot trial: should significance thresholds be reconsidered? *BMC Medical Research Methodology*, *14*(1), 41-49. https://doi.org/10.1186/1471-2288-14-41

Main, S., O'Rourke, J., Morris, J., & Dunjey, H. (2016). Focus on the journey, not the destination: Digital games and students with disability. *Issues in Educational Research*, *26*(2), 315-331.

Montgomery, C., & Smith, L. (2015). Bridging the gap between research and practitioners. *Die Unterrichtspraxis/Teaching German*, *48*(1), 100-113. https://doi.org/10.1111/tger.10183

Paulic, D. (2017). Pilot trial of a knowledge sharing intervention program with digital games. (Doctoral dissertation, University of Phoenix). Retrieved from https://eric.ed.gov/?id=ED586859

Petri, G., & Gress von Wangenheim, C. (2016). How to evaluate educational games: A systematic literature review. *Journal of Universal Computer Science*, *22*(7): 992-1021.

Petri, G., & Gress von Wangenheim, C. (2017). How games for computing education are evaluated? A systematic literature review. *Computers & Education*, *107*, 68-90. https://doi.org/10.1016/j.compedu.2017.01.004

Pritchard, K., & Whiting, R. (2012). Autopilot? A reflexive review of the piloting process in qualitative e-research. *Qualitative Research in Organizations and Management,* *7*(3), 338-353. https://doi.org/10.1108/17465641211279798

Procci, K., Lakhmani, S., Hussain, T., & Bowers, C. (2014). Opening cinematics: Their cost-effectiveness in serious games. *Simulation & Gaming*, *45*(1), 93-124. https://doi.org/10.1177/1046878113508515

Rajasekar, A., & Kumar, V. (2019). Randomized controlled trials: Gold standard of evidence. *Drug Invention Today*, *12(*6), 1215-1217.

Randel, B., & Clark, T. (2013). Measuring classroom assessment practices. In J.H. McMillan *SAGE handbook of research on classroom assessment* (pp. 145-163). Thousand Oaks, CA: Sage Publications, Inc. https://doi.org/10.4135/9781452218649.n9

Read, J., & Shortell, S. (2011). Interactive games to promote behavior change in prevention and treatment. *Journal of the American Medical Association*, *305*(16), 1704-1705. https://doi.org/10.1001/jama.2011.408

Reeves, B., Cummings, J., Scarborough, J., & Yeykelis, L. (2015). Increasing energy efficiency with entertainment media: An experimental and field test of the influence of a social game on performance of energy behaviors. *Environment and Behavior*, *47*(1), 102-115. https://doi.org/10.1177/0013916513506442

Richard V., & Belanger, M. (2018). Accepting research: Teachers' representations of participation in educational research projects. *International Journal of Educational Methodology*, *4*(2), 61-73. https://doi.org/10.12973/ijem.4.2.61

Schneider, S., Schmidli, H., & Friede, T. (2013). Blinded and unblinded internal pilot study designs for clinical trials with count data. *Biometrical Journal*, *55*(4), 617-633. https://doi.org/10.1002/bimj.201200189

Serret, S., Hun, S., Lakimova, G., Lozada, J., Anastassova, M., Santos, A. ... Askenazy, F. (2014). Facing the challenge of teaching emotions to individuals with low- and high-functioning autism using a new serious game: a pilot study. *Molecular Autism*, *5*(1), 1-35. https://doi.org/10.1186/2040-2392-5-37

Sims, W., Riley-Tillman, C., & Cohen, D. (2017). Formative Assessment using direct behavior ratings: Evaluating intervention effects of daily behavior report cards. *Assessment for Effective Intervention*, *43*(1), 6-20. https://doi.org/101177/1534508417708183

Smith, P., Morrow, R., & Ross, D. (Ed.). (2015). *Preliminary studies and pilot testing. In field trials of health interventions: A toolbox* (3rd ed.). Oxford, UK: OUP Oxford.

Stouck, J. (2019). Complementary learning: Piloting a blended format for Canadian composition courses. *Canadian Journal for the Scholarship of Teaching & Learning*, *10*(1), 5-20. https://doi.org/10.5206/cjsotl-rcacea.2019.1.7992

Strader, T. (2011). *Digital product management, technology, and practice: Interdisciplinary perspectives*. New York, NY: Business Science Reference. https://doi.org/10.4018/978-1-61692-877-3

Sullivan, G., & Sargeant, J. (2011). Qualities of qualitative research: Part I. *Journal of Graduate Medical Education*, *3*(4), 449-452. https://doi.org/10.4300/JGME-D-11-00221.1

Toomey, M. (2017). Engaging the enemy: Computer games in the English classroom. *Literacy learning: The Middle Years*, *25*(3), 38-49.

Tuzun, H., Barab, S., & Thomas, M. (2019). Reconsidering the motivation of learners in educational computer game contexts. *Turkish Journal of Education*, *8*(2), 129-159. https://doi.org/10.19128/turje.546283

Vate-U-Lan, P. (2015). Transforming classrooms through game-based learning: A feasibility study in a developing country. *International Journal of Game-Based Learning*, *5*(1), 46-57. https://doi.org/10.4018/ijgbl.2015010104.

Whitehead, A., Sully, B., & Campbell, M. (2014). Pilot and feasibility studies: Is there a difference from each other and from a randomized controlled trial? *Contemporary Clinical Trials*, *38*(1), 130-133. https://doi.org/10.1016/j.cct.2014.04.001

Wiens, P. (2014). Using a participant pool to gather data in a teacher education program: The course of one's school's efforts. *Issues in Teacher Education*, *23*(1), 177-206.

Wiseman, A., Whitty, G., Tobin, J., & Tsui, A. (2010). The Uses of Evidence for Educational Policymaking: Global Contexts and International Trends. *Review of Research in Education*, *34*(1), 1-24. https://doi.org/10.3102/0091732X09350472

Wolfe, B. (2013). The value of pilot studies in clinical research: a clinical translation of the research article titled "In search of an adult attachment stress provocation to measure effect on the oxytocin system." *Journal of the American Psychiatric Nurses Association*, *19*(4), 192-194. https://doi.org/101177/1078390313493248

Wood, C., Mentzelopolous, M., & Protopsaltis, A. (2015). EdCCDroid: An education pilot prototype for introducing Code-Combat using LUA. In D. Preuveneers (Ed.), *Workshop Proceedings of the 11th International Conference on Intelligent Environments* (pp. 353-340). Prague, Czech Republic: University of Westminster.