# Usability in Blended Learning Programs within Health Professions Education:

# A Scoping Review

Anish Arora

Department of Family Medicine

McGill University, Montreal

August 2019

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science in Family Medicine with a specialization in Medical Education

## Abstract

**Background:** Blended learning, an educational pedagogy that combines both traditional face-to-face teaching methods with technological innovations, is being implemented in training programs worldwide and specifically within the field of health professions education. Research indicates that usability, a multidimensional concept used to evaluate products and services, can be of paramount importance when evaluating the effectiveness, efficiency, and satisfaction with the use of blended learning programs (BLPs). Despite its importance, knowledge about the concept of usability and its consideration in BLPs within the field of health professions education remains lacking.

**Objective:** The purpose of this study is to develop a foundational understanding of how BLPs have generally been and are currently being evaluated in the field of health professions education in relation to the concept of usability and its primary components (effectiveness, efficiency, and satisfaction).

**Methods:** A scoping review guided by the framework outlined by Arksey & O'Malley, 2005 was conducted. The PRISMA-ScR checklist illustrated by Tricco et al., 2018 was also corroborated when completing this study. The databases searched were Scopus and ERIC (EBSCO). Searches were carried out in an iterative manner, where new terms were identified and added to the search between September 17, 2018 and September 21, 2018. Screening was conducted via the use of a questionnaire guide and in significant collaboration with academic liaison librarians and co-authors. Charted data was validated by a co-author. Thematic analysis was then conducted by two independent reviewers.

**Results:** The search strategy identified 8626 studies of potential relevance for the scoping review. After title and abstract screening, 508 studies were identified to be full-text reviewed. Full-text

review yielded 53 studies to be included in this scoping review. No study evaluated the overall concept of usability. 47/53 studies utilized a survey/questionnaire or feedback tool to evaluate their BLP. 33/47 studies did not indicate if their tool was reliability tested, standardized, or validated. Qualitative content analysis indicated that scholars do evaluate for the usability components *effectiveness, efficiency, and satisfaction*. Thematic analysis found three overarching themes: (1) Avoiding the 'Usability' Label and Using Undefined Related Terms such as *Helpfulness* and *Usefulness*; (2) Confusing Conceptualization of the Components of Usability (i.e. Effectiveness, Efficiency, & Satisfaction); and (3) Lack of Consensual Approach to Evaluation. These themes ascertained that the concept of usability is discussed across studies that met eligibility criteria, albeit implicitly. Authors were found to use multiple different terms to describe the same concepts. Where the same terms are applied across studies, different connotations are found to be applied to these terms. However, 31 key concepts were identified as associated with usability components (effectiveness, efficiency, and satisfaction), and two related concepts (accessibility and user experiences). Seven key "take home" ideas of what BLP evaluations should include were also identified through thematic analysis.

**Conclusion:** A lack in the conceptual understanding of usability and its associated terms is found across the literature. The results suggest that there is no consensus among researchers regarding evaluation terminology or methodology in this context. This review presents several key findings that help to establish a fundamental understanding of BLP evaluations with respect to usability in the field of health professions education. The findings of this study will be of paramount importance to stakeholders that plan on conducting evaluations of BLPs or developing new tools or frameworks to evaluate BLPs using the concept of usability in the context of health professions education. These findings will be of particular interest to family medicine education research as

this relatively underdeveloped sub-field of inquiry continues to advance and as family physicians continue to demonstrate their paramount importance in ensuring health system robustness amidst an ever-advancing technological era.

## Résumé

**Contexte:** L'apprentissage mixte, pédagogie combinant à la fois des méthodes d'enseignement traditionnelles en face-à-face et des innovations technologiques, est actuellement mis en œuvre dans le cadre de programmes de formation dans le domaine de la formation aux professions de la santé. Les recherches indiquent que la facilité d'utilisation, un concept multidimensionnel utilisé pour évaluer les produits et services, peut revêtir une importance primordiale pour évaluer l'efficacité, l'efficience et la satisfaction liées à l'utilisation de programmes d'apprentissage mixte (BLP). Malgré son importance, les connaissances sur le concept de convivialité et sa prise en compte dans les BLP dans le domaine de la formation des professionnels de la santé font encore défaut.

**Objectif:** Le but de cette étude est de développer une compréhension fondamentale de la façon dont les BLP ont été et sont actuellement évalués dans le domaine de la formation des professionnels de la santé en relation avec le concept de convivialité et ses composantes principales (efficacité, efficience et satisfaction).

**Méthodes:** Une étude de cadrage a été réalisée, guidée par le cadre défini par Arksey & O'Malley, 2005. La liste de contrôle PRISMA-ScR illustrée par Tricco et al., 2018, a également été corroborée lors de la réalisation de cette étude. Les bases de données consultées étaient Scopus et ERIC (EBSCO). Les recherches ont été effectuées de manière itérative. De nouveaux termes ont été identifiés et ajoutés à la recherche entre le 17 et le 21 septembre 2018. Ils ont été effectués à l'aide d'un guide et en collaboration étroite avec des bibliothécaires. Les données cartographiées ont été validées par un co-auteur. Une analyse thématique a ensuite été menée par deux examinateurs indépendants.

**Résultats:** La stratégie de recherche a identifié 8626 études potentiellement pertinentes pour l'examen de la portée. Après sélection du titre et du résumé, 508 études ont été identifiées comme faisant l'objet d'une révision en texte intégral. L'examen en texte intégral a permis d'inclure 53 études dans cet examen. Aucune étude n'a évalué le concept général de convivialité. 47/53 études ont utilisé une enquête / un questionnaire ou un outil de feedback pour évaluer leur BLP. 33/47 études n'ont pas indiqué si leur outil avait été testé, normalisé ou validé en termes de fiabilité. L'analyse qualitative du contenu a montré que les spécialistes évaluent l'efficacité, l'efficience et la satisfaction des composants liés à la convivialité. L'analyse thématique a permis de dégager trois grands thèmes: (1) éviter l'étiquette utilisabilité et utiliser des termes associés non définis, tels qu'utilité; (2) Conceptualisation déroutante des composantes de la convivialité (c.-à-d. Efficacité, efficience et satisfaction); et (3) le manque d'approche consensuelle en matière d'évaluation. Ces thèmes ont permis d'établir que le concept de convivialité fait l'objet d'une discussion entre les études répondant aux critères d'éligibilité, même implicitement. On a constaté que les auteurs utilisaient plusieurs termes différents pour décrire les mêmes concepts. Lorsque les mêmes termes sont appliqués à travers les études, différentes connotations s'appliquent à ces termes. Cependant, 31 concepts clés ont été identifiés comme étant associés à des composants de convivialité (efficacité, efficience et satisfaction) et à deux concepts connexes (accessibilité et expériences utilisateur). Sept idées clés à emporter de ce que les évaluations du BLP devraient inclure ont également été identifiées à travers une analyse thématique.

**Conclusion:** La compréhension conceptuelle de la convivialité et de ses termes associés fait défaut dans la littérature. Les résultats suggèrent qu'il n'y a pas de consensus parmi les chercheurs concernant la terminologie ou la méthodologie d'évaluation dans ce contexte. Cette revue présente plusieurs résultats clés qui aident à établir une compréhension fondamentale des évaluations du

BLP en ce qui concerne la convivialité dans le domaine de la formation des professionnels de la santé. Les conclusions de cette étude revêtiront une importance capitale pour les parties prenantes qui prévoient de mener des évaluations des points de vente blancs ou de développer de nouveaux outils ou cadres pour évaluer ces points de vue en utilisant le concept de convivialité dans le contexte de la formation des professionnels de la santé. Ces résultats seront particulièrement intéressants pour la recherche en médecine familiale, car ce sous-domaine de recherche relativement sous-développé continue de progresser et que les médecins de famille continuent de démontrer leur importance primordiale pour assurer la robustesse du système de santé dans une ère technologique en constante évolution.

## Table of Contents

Abstract ..... 2
Résumé ..... 5
List of Abbreviations ..... 10
List of Figures ..... 11
List of Tables ..... 12
Acknowledgements ..... 13
Preface Regarding Thesis Format ..... 15
Contribution to Original Research ..... 16
Contribution of Authors ..... 18
1. Introduction ..... 19
2. Literature Review ..... 23
2.1 The Evolution of Learning Pedagogies through the Introduction of New Technologies ..... 23
2.2 The Need to Comprehensively Evaluate Blended Learning Programs & Existing Evaluative Frameworks ..... 26
2.3 Usability as a Key Concept in the Evaluation of Blended Learning Programs ..... 28
2.4 Components & Methods to Consider When Evaluating for Usability ..... 30
2.5 Summary of the Literature Review & Rationale for this Study ..... 32
3. Methodology ..... 33
3.1 Scoping Review ..... 33
3.2 Academic Liaison Librarian Involvement ..... 34
3.3 Step 1: Identify the Research Question ..... 34
3.4 Step 2: Identifying Relevant Studies ..... 34
3.5 Step 3: Study Selection ..... 35
3.6 Step 4: Charting the Data ..... 38
3.7 Step 5: Collating, Summarizing, and Reporting the Results ..... 39
4. Results ..... 42
4.1 Studies that met Eligibility Criteria & Charted Data ..... 42
4.2 Quantitative Analysis of Extracted Data ..... 62
4.3 Qualitative Thematic Analysis ..... 69
Part 1: Inductive Analysis ..... 70
Theme 1: Avoiding the ‘Usability’ Label and Using Undefined Related Terms ..... 70

Theme 2: Confusing Conceptualization of the Components of Usability (i.e. Effectiveness, Efficiency, & Satisfaction) ..... 74
Theme 3: Lack of Consensual Approach to Evaluation ..... 79
Part 2: Deductive Analysis ..... 82
4.4 Summary of Findings from Qualitative Thematic Analysis ..... 88
5. Discussion ..... 89
Limitations ..... 94
Strengths ..... 95
Contributions and Future Directives ..... 96
Contributions Specific to Family Medicine ..... 97
6. Conclusion ..... 99
References ..... 100
Appendix ..... 118
Appendix 1: Search Strategy for Scopus ..... 118
Appendix 2: Search Strategy for ERIC (EBSCO) ..... 120
Appendix 3: Title, Abstract, & Full-Text Screening Questionnaire ..... 122
Appendix 4: NVIVO Codes for Qualitative Analysis ..... 124
Appendix 5: PRISMA-ScR Checklist ..... 128
Appendix 6: Institutional Research Ethics Board Approval ..... 129
Appendix 7: Institutional Research Ethics Board – Continuing Review Acceptance ..... 130

## List of Abbreviations

BL = Blended Learning

BLP = Blended Learning Program

BLPs = Blended Learning Programs

FC = Flipped Classroom

ISO = International Organization for Standardization

LMS = Learning Management System

LMSs = Learning Management Systems

## List of Figures

1. PRISMA flow diagram of included/excluded studies…….....…Page 43

2. The number of studies that met eligibility criteria by year of publication……Page 62

3. The frequency of labels used to refer to “blended learning” by year……Page 63

4. Total number of studies that met eligibility criteria from each country/region and the labels used to refer to blended learning by each respective country/region……...…….Page 64

5. Frequency of usability components being evaluated by year……...…Page 65

6. Visualization of the frequency of studies by their specific population cohorts……Page 66

7. Concept Map Developed from the Deductive Findings of the Thematic Analysis……..Page 91

## List of Tables

1. ISO Definitions of Usability Components and Examples of Coding…….....…Page 40

2. Synthesized Table of Extracted Data for Studies that Met Eligibility Criteria…….Page 44

## Acknowledgements

Bahá'u'lláh, the founder of the Bahá'i faith, asks us to "regard man as a mine rich in gems of inestimable value. Education can, alone, cause it to reveal its treasures, and enable mankind to benefit there-from." To me, this means that humanity has incredible potential to create positive change, but this potential can only be transformed into action through polishing, or more precisely, through educating individuals.

For me, there are several people that have significantly assisted in polishing my potential and allowing for me to produce this thesis.

First and foremost is Dr. Charo Rodriguez, my primary supervisor. Charo, thank you for always being there for me – as a supervisor, mentor, and most importantly, as a strong support system. I am truly grateful for all the advice and guidance that you have provided me with, in both writing my thesis and in my overall life. Secondly, I want to thank my co-supervisor, Dr. Tibor Schuster. Tibor, thank you for always having your door open to me and for always having my back during these last two years. Charo and Tibor, you both have been and continue to be instrumental in my growth as a professional and individual. I am so grateful to you both.

I also want to say thanks to Dr. Tamara Carver, the third individual on my thesis committee. Thank you, Tamara, for being such a cool person, for providing me with ample support in relation to my thesis, and for just having super fun conversations about random life things with me over these last two years.

Similarly, thank you Sherrie Child for assisting me on the administrative end and for being such a cool person. Your stories and jokes never cease to put a smile on my face.

Thank you, Genevieve Gore, Ekatarina Grguric, Lucy Kiester, and Vera Granikov. Without the help of you four librarians, I would not have been able to conduct such an interesting project.

I also want to thank Dr. Howard Bergman and the McGill Department of Family Medicine Graduate Research Program for financially supporting my first year at McGill, as well as the Social Science and Humanities Research Council for supporting my second year through a CGS-M scholarship. This financial support was key to me being able to focus my efforts on completing this thesis and growing as a strong researcher.

Most importantly, I want to thank my family: Usha, Mehesh, and Rohit Arora. Without your support, all of this would not have been possible. Even though I haven't been able to visit home as often as I had hoped, your daily calls made it feel like I really wasn't that far away. Thank you, mom and dad, for shaping who I am today. Thank you for showing me the joy of working hard against all odds. And thank you for always keeping me strong.

I sincerely hope that I have made all of you proud.

## Preface Regarding Thesis Format

This thesis has been written in a traditional format.

## Contribution to Original Research

To my knowledge, this is the first study that comprehensively searches the literature for evaluation studies in the field of health professions education, with the use of a strong definition for blended learning and usability. This study presents several major results, the most important of which is that the concept of usability has yet to be explicitly utilized to evaluate blended learning programs, although all studies implicitly discuss the concept of usability and do evaluate for one or more components of usability (i.e. effectiveness, efficiency, and satisfaction).

The results from the quantitative analysis portion of this study contribute towards developing an understanding of the overall extent and range of the blended learning program evaluations in the field of health professions education. For instance, a key finding in this section was that, in most studies, blended learning program evaluations take place after the program has been completed and through questionnaires/surveys or feedback tools. Often no indication regarding instrument validation, standardization, or reliability-testing regarding these survey instruments is provided in these studies.

Through qualitative content analysis it was identified that scholars do evaluate for the usability components of effectiveness, efficiency, and satisfaction. Following this, the three themes found in the qualitative inductive semantic thematic analysis section of this study assist in developing an understanding of: (1) the different terms that authors apply across studies to define similar concepts; (2) the different connotations that authors ascribe to similar terms and concepts across studies; and (3) the differences in the evaluation methods undertaken by authors across the literature. The deductive analysis that followed the inductive analysis section assisted in understanding how the International Organization for Standardization's definitions of usability, its components, and its related concepts bring clarity and consistency to the terms, concepts, and

evaluations presented across the included literature. This section further gave rise to a concept map and seven key ideas of what future blended learning program evaluations should consider. These two outcomes in particular of the qualitative analysis will be of paramount importance for stakeholders that plan on conducting evaluations of blended learning programs or developing new frameworks to evaluate BLPs using the concept of usability.

Therefore, the findings of this study have the potential to directly impact BLP evaluations across the field of health professions education, which in turn, has the potential to strengthen healthcare delivery across various levels of care.

## Contribution of Authors

I completed this thesis under the guidance and supervision of my primary supervisor, Dr. Charo Rodriguez (ChR); my co-supervisor, Dr. Tibor Schuster (TS); and a member of my Thesis Advisory Committee, Dr. Tamara Carver (TC). These three individuals assisted in the planning phase of the thesis. Furthermore, these three individuals met with me on a regular basis (every 2 to 3 weeks) to ensure that I was completing all aspects of this thesis appropriately and accurately. TC was consulted significantly during the screening and data extraction sections of this study. TS guided the quantitative analysis portion of this study. ChR assisted significantly with the thematic analysis and the interpretation of the results. A fellow MSc student, Mr. Matthew Hacker-Tepper (MHT), was recruited to validate the charted data and to act as an independent secondary coder for the thematic analysis. Lastly, my entire Thesis Committee (ChR, TS, and TC) provided major edits and feedback to multiple drafts of this thesis.

## 1. Introduction

The purpose of this study is to map current knowledge about and develop a foundational understanding of how the concept of usability has been conceptualized and evaluated in the context of blended learning programs (henceforth BLPs) in the field of health professions education.

To keep up with the increasing pace of technological advancement, as well as maximize the benefits of this digitalized world, we need to transform our learning environments [1-3]. Blended learning (henceforth BL), an educational pedagogy that combines traditional face-to-face teaching strategies with technological platforms and innovations, provides this transformative potential for higher education [1-2]. Over the last decade, BLPs have been implemented to a high degree in educational interventions worldwide, and particularly within the field of health professions education [4-7].

Several studies indicate that BLPs are highly effective in providing opportunities for meaningful learning [1, 5, 8-14]. A potential explanation for this finding is that by adding technology into educational settings, learners (i.e. participants of an educational program such as undergraduate students or professionals taking part in a continuing professional development imitative) can tailor their learning experiences according to their needs and objectives, which in turn, enables them to control the content, sequence, pace, and time of their learning [5, 10]. Furthermore, along with providing learners with increased flexibility in their learning, BLPs provide a cost-saving potential for educational institutions in the long run [15].

BL has been defined in different ways in the literature. In general, any program that combines the use of traditional face-to-face teaching methods with online learning environments can be considered as a BLP [1, 2, 16, 17]. Such a broad definition allows for BLPs to be developed very differently across disciplines, programs, faculties, and institutions [17].

However, some authors conceive BL in a more restrictive manner [16, 17], which ultimately assist in establishing a distinction between BLPs, web-facilitated programs, and online learning programs [16]. These authors indicate that a truly BL model delivers approximately 30 to 79% of its educational material through online learning methods [16]. These online learning methods are traditionally asynchronous, meaning that learners can learn material at different times and in different locations [1, 17]. Web-facilitated learning programs, in comparison, only utilize online learning methods for 1 to 29% of the program, and consequently, these programs are often considered to be enhanced face-to-face courses [16]. On the other end of the spectrum, fully online courses are those that deliver their educational material through online means for more than 80% of the program, and often, these programs do not have any face-to-face meetings [16].

Recently, the utilization of a learning management system (henceforth LMS) has also become a key determinant to ensuring that an educational program is in fact implementing a truly blended model of education [4, 9, 18]. An LMS is a software application that allows for easy learner progress tracking, organization of course content, and online communication between learners and instructors [2, 4, 8, 9, 18]. Although LMSs have been predominantly used in e-learning settings, they are of significant benefit to BLPs [2, 4, 8, 18]. The rationale for this is that BLPs utilize a significant amount of web-based communication, harbor extensive amounts of content, and consist of a great deal of work for both learners and instructors [2, 4, 8, 9, 18]. Thus, a web-based LMS must be employed to organize and manage the content, facilitate efficient communication, and accurately track learner progress in BLP settings [2, 4, 6, 8, 9, 18].

The overall value of BLPs is that these programs are more than just a combination of traditional learning methods and online learning environments [1]. When implemented appropriately, BLPs allow for learners to feel in control of their learning, while also empowering

educators to effectively guide and monitor learner progress through LMSs as mentioned above [4, 18]. Again, this control in learning and effective learner tracking relies heavily upon the use of an LMS. In particular, these systems allow educators to accurately identify where learners are at in relation to the course content, as well as assist educators in identifying potential issues learners may have while progressing through the course [2, 4, 6, 8, 9, 18]. As educators can track learner progress on their own time, students may not always be aware that their progress is being monitored [2, 4, 6, 8, 9, 18]. These advantages that LMSs, and in turn BLPs, bring to educational settings make BL an excellent innovative educational model for adult learning [2, 10, 16].

However, difficulties exist in adopting BL instruction into an educational system. Since the implementation of BLPs require an initial large investment in faculty training, time, and money [10], it appears necessary to comprehensively and rigorously evaluate both the technological platforms and the in-class educational aspects of BLPs prior to their widespread implementation [1, 10]. Here, *usability* appears to be one of the most important dimensions of the BLPs that needs to be considered and evaluated for [19, 20, 26].

Usability is a multidimensional concept defined by the International Organization for Standardization (henceforth ISO) as the "extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" [21]. According to this definition, usability is composed of three primary components, namely effectiveness, efficiency, and satisfaction [20-27]. To our knowledge, this is the most recently revised definition of the concept [21]. Furthermore, this definition has been developed in collaboration with several experts on this concept [21, 23, 25]. Also, this definition explicitly articulates the application of usability in both the technological and service aspects of a system [21]. This makes the application of the ISO definition ideal for the current context [21, 23,

25]. Through measuring for its components – often prior to the implementation of a program or technology – the concept of usability allows creators of BLPs to develop a stronger understanding of whether their program is well designed and well received by users [11, 19]; if the program has the potential to facilitate learning [15, 19]; and how the program can be improved for the future [28].

However, although usability is a highly researched and heavily defined concept in the field of integrated technology [29, 30], research on usability in BLPs has been poorly developed [19]. In fact, studies indicate that empirical evaluations of usability are rarely done in the field of education [22, 31, 32], and this is especially true in the context of medical education [19, 32]. This lack of translation of usability into the field of medical education, and more broadly health professions education, may be due to the controversy surrounding the definition of usability – as several different definitions exist – [19-27] or potentially because there is a lack of validated usability evaluation methods that consider users as learners [33]. It is important to note briefly that when speaking of health professions education, which is the focus of this thesis, reference is being made to educational initiative for both trainees (e.g. undergraduate medical students, nursing students, and physiotherapy students) and trained professionals (e.g. family doctors, nurses, physiotherapists taking part in continuing professional development of faculty development initiatives).

## 2. Literature Review

In this section, I will briefly discuss the emergence of new learning pedagogies in health professions education as technology advanced over the last two decades (i.e. how BLPs came to exist), along with the naming trends for these learning pedagogies. This will allow readers unfamiliar with BLPs to understand specific nomenclature, as well as provide both novices and experts in this field with a succinct understanding of the context in which education is currently delivered. Following this, I will discuss the need to comprehensively evaluate BLPs. Brief references will be made to currently existing evaluation frameworks. I will then highlight what is known about the multidimensional concept of usability and how it is generally evaluated for. Finally, I will end this section by summarizing the main points of the literature review and the rationale for this study.

### 2.1 The Evolution of Learning Pedagogies through the Introduction of New Technologies

Prior to the invention and adoption of the world wide web in the late 1980s-early 1990s, *multimedia learning* was an important buzzword in the field of health professions education [9, 10]. This learning pedagogy utilized two or more media, such as texts, graphics, animations, audio, or video, to develop content that engaged learners and could be accessed through a computer [9, 10]. At the time, the technology used to store these learning media were primarily floppy disks and CD-ROMs [9, 34]. Upon arrival of the internet, a shift towards the *electronic-learning* pedagogy, or *e-learning* for short, was seen – specifically between 1991 and 1998 [9, 10]. E-learning methods were generally asynchronous [9, 35] and had a variety of names ascribed to them such as web-based learning and internet-based learning [10]. Within the overall pedagogy of e-learning, there were two main sub-pedagogies: *distance learning* and *computer-assisted instruction*, the latter of which was also referred to by various terms including the following:

computer-assisted learning, computer-assisted teaching, computer-based learning, and computer-based training [10]. Distance-learning referred to e-learning that occurred remotely from the educational institution (i.e. learners can access learning content at home) [9, 10], whereas computer-assisted instruction essentially referred to the general use of computers to assist in delivering educational content [10].

When discussing e-learning in literature, authors indicate that it has significant benefits over traditional learning pedagogies which involve only face-to-face teaching [8, 35]. Educational platforms that utilized technology were said to have provided learners with the ability to tailor their learning experiences according to their needs and objectives [10]. In fact, when adding technology into educational interventions, learners gained the ability to control the content, sequence, pace, and time of their learning [5, 10]. Moreover, implementing information technology into educational systems begins bridging into the concept of individualized learning [10], which is understood as not only allowing students to control their learning but also enabling educators to adapt their teaching strategies to their students [36]. Individualized learning is critical in making learning empowering and learners successful [36].

A 2007 review of e-learning in medical education further emphasized the benefits of implementing technology within this field [8]. The review described how technology allows for a way to keep up with the increasingly massive amount of information that can be found, and at the same time, provide reliable and reusable content in a very convenient format for users [8]. By 2008, a comprehensive two-part guide on e-learning [11, 35] was developed which indicated that e-learning had become mainstream in most medical schools [35].

However, back tracking a few years, around the mid-2000s, an increasing shift towards *blended learning* (BL) was beginning to be seen worldwide [5, 6, 37]. BL, also referred to as

hybrid learning and hybrid teaching [38], was initially described as a platform of instruction that combines both face-to-face learning with the utilization of technology [1, 2]. This blend of learning pedagogies was discussed as providing learners with increased flexibility in their learning through the inclusion of technology, and concurrently, it retained the principles, values, and overall advantages of face-to-face instruction [5, 38, 39]. Moreover, BL models were frequently shown to be more effective and efficient than traditional classroom models of teaching [12-14] and were also discussed to have provided educational institutions a cost-saving potential in the long run [15]. These are possibly key reasons for why BL models are being implemented in educational interventions globally, and especially in the field of health professions education [4-7].

As technology has developed, more refined definitions and criteria for what constitutes a BL model of instruction have been developed [5]. One key definition provided by Allen et al. in 2007, indicates that a BLP is one that delivers approximately 30 to 79% of its educational material through generally asynchronous online learning methods [16]. This allow for learners to learn content at different times and in different locations [1, 17]. Allen et al. go on to describe that educational programs that utilize online learning methods for only 1 to 29% of the program are better considered to be *web-facilitated learning* programs [16]. The rationale for this is that these programs are essentially face-to-face teaching programs that incorporate a minimal amount of technology to enhance the learning process, but not necessarily to teach content [16]. On the other end of the spectrum are fully online courses which deliver their education material through online means for greater than 80% of the program [16]. These courses are also not considered to be truly blended because they are essentially online classes that utilize a very small amount of face-to-face contact to slightly enhance the learning process [16]. An example of this can be a learning program

that has one introductory face-to-face class in which the syllabus is explained to learners prior to them engaging with asynchronous content online for the rest of the term.

A specific sub-type of the BL pedagogy is the *flipped classroom* (henceforth FC) model, also referred to as flipped learning, flipped teaching, inverted learning, and inverted classroom [40]. The FC model was first discussed in the literature in the mid to late-2000s, and its development is generally accredited to two high school teachers in the United States, namely Aaron Sams and Jonathan Bergmann [41]. In this pedagogy, passive learning content is provided online and outside of class time, while active learning strategies such as discussions are utilized during class time to assist in consolidating the learning [42, 43]. The FC model has been especially well received in institutions of higher education and specifically in the field of health professions education [44, 45].

## 2.2 The Need to Comprehensively Evaluate Blended Learning Programs & Existing Evaluative Frameworks

To adopt technological platforms into educational systems, an initial large investment in faculty training, time, money, and space in many cases is needed [10]. Thus, studies express the need to critically evaluate them prior to their widespread implementation [1]. This is also the case in the field of health professions education [8, 10].

In the specific subfield of medical education (within health professions education), Cook & Ellaway recently discuss that the evaluation for learning programs which utilize information technology is often hindered for several reasons, including: (1) the lack of comprehensiveness of the evaluations; (2) a disconnect between what happened during the learning program and what was evaluated; and (3) the absence of a thorough framework or method for evaluation [46]. These authors further describe that although some medical education researchers have presented potential

guidelines and the use of specific instruments for evaluating the quality of computer-based learning models, none of these evaluative frameworks or instruments are thought to be comprehensive [46].

To fulfil this research gap in technology-enhanced education evaluation, Cook & Ellaway developed a new framework that references several previously developed evaluation frameworks such as the: Kirkpatrick Framework; CIPP Framework; and the Quality Matters Evaluation Program [46]. Cook & Ellaway outline seven broad areas of evaluation activity in their framework which include: (1) *conducting a needs analysis and environmental scan*; (2) *documenting processes, decisions, and the final product*; (3) *testing usability*; (4) *documenting key events during implantation and the final product*; (5) *assessing participant experiences and satisfaction;* (6) *assessing learning outcomes*; and (7) *estimating costs, reusability, and sustainability* [46]. The authors then indicate that completing all of these evaluations may be difficult to do and thus present a "minimal recipe for technology-enhanced learning evaluations" or in other words a "simple plan" which includes: (1) *performing usability testing*; (2) *documenting key elements of the final product*; (3) *administering instruments to capture the perceptions of both students and instructors*; and (4) *preparing and administering course-specific evaluations of Kirkpatrick level 2 outcomes (knowledge, skills, attitudes)* [46].

However, although incredibly comprehensive in its approach, their evaluative framework for technology-enhanced learning presents two major issues in relation to BLPs: (1) the framework is too broad: it was developed to address learning programs that utilize technology to various degrees, and thereby, is not built specifically to deal with the evaluation of BLPs; and (2) even the "simple" approach to evaluation using this framework may be too long and complex to conduct in a time sensitive and low resource setting [46].

Interestingly, in 2017, Chmiel et al. addressed this need for a comprehensive framework to evaluate BLPs in the field of health professions education by developing their own framework [47]. Their well-developed mixed methods evaluative framework not only evaluates both the distance and face-to-face aspects of a BLP, but it also outlines the need to and process of evaluating all agents involved in a BLP (i.e. learners, faculty, and the administration team) [47]. However, Chmiel et al. clearly indicate that their evaluative framework is resource intensive and that further steps need to be taken to ensure the development of a tool that balances comprehensiveness and efficiency [47].

In sum, although efforts have been made to structure BLP evaluations, this initial literature review has found no evaluative frameworks that allowed for a fairly-comprehensive evaluation, balancing resources and time, of BLPs in the field of health professions education.

Interestingly, both evaluation frameworks mention usability testing and briefly highlight its importance. However, these authors' conceptualization of the term is rather restrictive since they only discuss evaluation of usability in relation to the technological aspects of the learning pedagogy.

## 2.3 Usability as a Key Concept in the Evaluation of Blended Learning Programs

Usability testing is a highly developed area of research and practice in the field of integrated technology [19]. It is simply defined as the ease with which something can be used [20]. The International Organization for Standardization (henceforth ISO), however, indicates that usability is a significantly more complex concept than just its commonly understood "ease-of-use" or "user-friendliness" definitions [21]. The ISO defines usability as the "extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" [21]. ISO discusses usability as an

outcome of interacting with a system, product, or service, rather than just as an attribute of a product [21]. Moreover, ISO lists several instances in which the concept of usability is highly relevant, including learning [21].

Literature has highlighted that the multidimensional concept of usability could be incredibly valuable in evaluating learning pedagogies that incorporate the use of technology [11, 19]. In fact, some scholars have identified usability of the user interface within a BLP as the primary criteria that stakeholders need to consider when attempting to create and implement effective blended learning programs [15].

It is important to note that although several other definitions of usability exist [20, 23, 48-51], I have adopted in this thesis the one that was developed by the ISO. There are several reasons for this. First, this definition was revised in 2018 based on two studies presented by Bevan et al. in 2015 and 2016 [23, 25], making the ISO definition the most up-to-date definition for usability to our knowledge. Secondly, three key revisions were implemented in the current definition of usability which include the following: (a) the words *system* and *service* were added, allowing for the scope of the definition to be extended towards evaluation for both technology and human-provided services; (b) the word *efficiency* has been redefined; and (c) the word *satisfaction* was clarified and discussed as applicable to a broader range of issues [21]. Finally, and importantly, the ISO definition clearly indicates that usability can be used to evaluate both the technological and face-to-face components of learning pedagogies, making it a comprehensive concept to ground an evaluation for BLPs in [21].

Although research has highlighted the importance of evaluating usability in educational interventions that utilize technology [19, 20], it has however rarely been done [22, 28, 31]. This is especially true in the context of health professions education [19, 32]. In fact, there is a growing

concern of technology not being able to impact learning as much as it is thought to [46, 52]; some have argued that this is possibly due to the lack of consideration regarding the usability of the technological platforms utilized in educational interventions [20, 31].

Within recent evaluative research in health professions education, only a handful of studies were found to consider usability, and these came more particularly from the sub-set of medical education research. Some of these works focus on the usability of e-learning platforms created for learners in medical school [53], and more specifically, for learners specializing in surgery [54], internal medicine and anesthesiology [55]. Furthermore, in these studies, usability evaluations were primarily conducted through questionnaires [53-55]. Only one of these studies utilized recordings of user interactions with the technology alongside a questionnaire to evaluate for usability [55].

In the few studies in the field of health professions education that evaluate educational initiatives for more than just medical education initiatives (i.e. interdisciplinary health education initiatives), usability was also found to be measured through questionnaires [56]. Some scholars also discuss the difficulties of usability when implementing a BLP [57]. However, their focus is predominantly on discussing the issues that creators of the program faced such as which theoretical approaches and design elements they should consider when creating their BLP, rather than on evaluating the BLP through the lens of the learners taking part in the program [57].

## 2.4 Components & Methods to Consider When Evaluating for Usability

Although very few studies exist that evaluate for usability regarding BLPs in the field of health professions education [22, 28, 31], a great deal of research in other fields of inquiry – specifically integrated technology – has been found which describes the methods for how usability

could be evaluated [19, 20, 24, 33, 62]. This section summarizes some significant knowledge from the literature regarding methods of usability testing.

Firstly, and in keeping with the ISO definition, studies indicate that usability should be evaluated with user responses towards three domains: *effectiveness*, how well users are able to achieve goals using the technological platforms; *efficiency*, how long and how much effort did it take to accomplish tasks and learn material using the technological platforms; and lastly, *satisfaction*, what is the comfort and level of acceptance towards the technological platforms used by learners [20, 22-24].

Next, when designing educational interventions that utilize technology, which are deemed as efficient and lead to satisfaction by users, three factors seem to be considered: the characteristics of the learner, the technology being utilized, and the context in which learning using the technology takes place [20, 33, 63].

Continuing, research indicates that 85% – 95% of issues with usability such as the lack of ease with navigating through websites or other online applications can be identified with feedback from approximately five to six users [19, 20, 62]. It is important to note that these usability evaluations primarily take place prior to the deployment of the technological program [19, 20, 62]. Furthermore, studies indicate that application of either questionnaires or through asking users to think aloud the process they would undergo to navigate through hypothetical scenarios with technology, are both valid methods of evaluating usability [20].

Some scholars have indicated that two out of the following three methods – user testing, semi-structured interviews, and surveys – need to be conducted in conjunction to comprehensively evaluate for usability [58]. However, no study was found to evaluate usability using qualitative methods such as semi-structured interviews and focus groups, even though semi-structured

interviews have been identified as an effective means for evaluating usability when conducted in combination with other methods [58] and focus groups have been discussed as being a key method for collecting data of the perceptions of users regarding services [59]. By relying heavily on questionnaires as a method of collecting data, we lack critical pieces of knowledge [58] that can be found through semi-structured interviews [60] and focus groups [61].

## 2.5 Summary of the Literature Review & Rationale for this Study

To summarize, we know that BLPs are being applied widely and that they are ubiquitous in the field of health professions education. This implies an increasing need for them to be comprehensively evaluated for. However, much still needs to be done to properly develop BLP evaluative research. The lack of development in this field is in part due to the lack of evaluation models that are specifically built to evaluate this type of learning pedagogy. I posit that usability could be instrumental in structuring a BLP evaluation model. Clear definitions regarding usability and its key components are provided by the ISO, and a rich body of knowledge exists on how usability could be evaluated for in disciplines such as information technology. However, despite its importance, the concept of usability has been poorly adopted and examined in the field of health professions education.

This study has therefore been conceived to develop a foundational understanding of how BLPs in the field of health professions education have been evaluated for in relation to the concept of usability, and more specifically, its primary components (i.e. effectiveness, efficiency, and satisfaction). This will not only address the knowledge gap present (i.e. the lack of understanding of how usability has been applied in this field of inquiry) but will also assist in culminating the disparate body of knowledge and evidence related to BLP evaluations in the present context, which in turn will identify and guide major areas of research related to BLP evaluations in the future.

# 3. Methodology

## 3.1 Scoping Review

A scoping review study was conducted. As the purpose of this study is quite broad and exploratory – to develop and essentially map out a foundational understanding of how BLPs have been and currently are evaluated in the field of health professions education in relation to the context of usability, or more specifically its primary components (i.e. effectiveness, efficiency, and satisfaction) – it will benefit highly from the application of a scoping review methodology [64-68].

Furthermore, the objective of this study is in agreement with the four reasons for why a scoping review methodology may be applied as discussed by Arksey & O'Malley, which are: (1) to provide an examination of the extent and nature of a specific body of research activity; (2) to determine if there would be value in conducting a systematic review; (3) to provide a summary of research findings and to disseminate this information; and (4) to identify if there are gaps in the existing body of literature [64]. Moreover, other authors also indicate that scoping reviews are utilized when the research question is broad, and the purpose of the investigation is exploratory [66-68].

This study was guided by the five-stage framework outlined by Arksey & O'Malley, 2005 [64]. The stages include: (1) identifying the research question; (2) identifying relevant studies; (3) study selection; (4) charting the data; and (5) collating, summarizing and reporting the results [64]. The PRISMA-ScR checklist illustrated by Tricco et al., 2018 was also corroborated when completing this study [65].

It is important to note that this review methodology does not seek to evaluate the quality of evidence, meaning that this review does not function to determine if the included studies have provided generalizable findings through a robust research process [64]. These tasks are more akin to those of the systematic review methodology [64]. Rather, the scoping review methodology, simply put, is used to better understand the content and context of a body of literature and assist in mapping out this understanding [64-68].

## 3.2 Academic Liaison Librarian Involvement

The review questions, search strategy, inclusion/exclusion criteria, and limits pertaining to this scoping review have been developed in close collaboration with three academic liaison librarians at McGill University, where: one is an expert in usability (EG); a second is an expert in conducting literature searches with the concept of medical and health professions education (LK); and the third is an expert in conducting literature searches with the concept of family medicine (GG). A fourth librarian associated with the Department of Family Medicine at McGill assisted in guiding the scoping review process (VG).

## 3.3 Step 1: Identify the Research Question

The following research question was used to guide this review: how has the concept of usability, or more precisely, its primary components (i.e. effectiveness, efficiency, and satisfaction) been defined and evaluated for in blended learning programs within the field of health professions education?

## 3.4 Step 2: Identifying Relevant Studies

This scoping review utilized two databases, namely Scopus and ERIC (EBSCO). Scopus was utilized as it is one of the largest global interdisciplinary databases which retains studies

conducted in the field of health professions education and technology [69]. ERIC (EBSCO) was utilized as it is one of the largest education databases in the world [70] which made it a logical secondary database to search since two of the main concepts in this study are “blended learning” and “health professions education.”

Numerous studies pertaining to effectiveness and satisfaction in relation to educational pedagogies exist in the literature. Furthermore, there are a multitude of well-defined terms that can be used to identify studies that fall under the concepts being explored in this study. Under the specific guidance of two librarians assisting with the development of the search strategy, it was decided that the use of Scopus and ERIC (EBSCO) would be enough for meeting the objective of this study. Hence, no grey literature was searched for in this study.

The search strategy for each database was comprehensive. Slight differences between the two searches can be seen due to variations in controlled vocabulary between the two databases. The strategy was focused on three main concepts: usability, blended learning, and health professions education. The full search strategy for both databases can be seen below in Appendix 1 and 2. Searches were carried out in an iterative manner where new terms were identified and added to the search between September 17, 2018 and September 21, 2018.

## 3.5 Step 3: Study Selection

All articles were imported into EndNote X8. All titles and abstracts were screened by me, but in close consultation with academic liaison librarians and co-authors. Consultation in this case referred to discussions and review sessions which functioned to assist in clarifying which studies met the inclusion and exclusion criteria. A questionnaire guide was created with respect to the inclusion and exclusion criteria (Appendix 3) and was utilized as another means of ensuring that minimal bias was present when screening. Items in this questionnaire guide had three options

"Yes," "No," and "Maybe." As titles were often ambiguous, the "Maybe" option allowed for many abstracts to be screened. Full-text screening was also conducted by me, again, in significant consultation with the members of my Thesis Advisory Committee. After all studies were screened initially, TC (Thesis Advisor) validated 36% of included studies to ensure that they matched the eligibility criteria. The limits and eligibility criteria for study inclusion and exclusion can be found below.

### *3.5.1 Inclusion Criteria*

Several inclusion criteria were agreed on for retrieving pertinent works. First, included studies had to empirically evaluate (using qualitative, quantitative, or mixed methods approaches) the blended learning delivery modality of health professions education programs in relation to either the overall concept of usability or one of its following major components: effectiveness, efficiency, satisfaction.

Second, studies had to describe the synchronous and asynchronous components of the BLP being evaluated or provide some indication about the number of hours of learning that each component of the BLP took. BLPs must have utilized online asynchronous technology for 30-79% of the educational delivery. This technology must have been accessible outside of the typical learning/teaching environment (i.e. at home). BLPs must have utilized a learning management system or indicate that learner use of online material was tracked.

Third, studies had to come from the field of health professions education (i.e. undergraduate or graduate education provided to students training to become health professionals, or continuing education/faculty development training provided to practicing health professionals [71]. A health professional can be considered any individual that functions to "maintain health in humans through the application of the principles and procedures of evidence-based medicine and

caring" [72]. In this context, only studies on those health professionals (students and practicing professionals – henceforth learners) that directly provide care to patients in some way (e.g. nurses, physicians, physiotherapists, pharmacists) were included (refer to the exclusion criteria below for further clarification). That said, learners (i.e. both students and practicing professionals taking part in an educational initiative) had to have been the primary individuals providing evaluations regarding the program (refer to the exclusion criteria below for further clarification).

### *3.5.2 Exclusion Criteria*

As included BLPs must utilize *online* asynchronous learning modalities for 30-79% outside of the classroom/clinical learning environment, studies using CD-ROMs, DVDs, and other downloadable software as their primary mode of asynchronous delivery were excluded. Similarly, studies that utilize simulation centres or computer labs as their primary technological component were also excluded as this technology is not accessible outside the general learning environment. Studies that pertain to learners (i.e. health professionals or students studying to become health professionals), that do not directly provide some form of care to human patients (i.e. veterinarians) were excluded as well. Studies that evaluated the BLPs of undergraduate courses that are not primarily delivered to health professional learners were also excluded – meaning that courses such as physiology or introductory psychology that are delivered by a faculty of science to all first-year learners, not necessarily health professional learners, were excluded from this study. Lastly, studies that solely conducted evaluations by faculty members (i.e. the instructors of a BLP and not the learners taking part in the BLP) were not retained in this study.

### *3.5.3 Limits*

Studies had to be written in English, come from a peer-reviewed journal article (as the terms utilized in this study are well-defined and fairly-well indexed), and be published and indexed between August 6, 1991 (as this was when the world wide web went live) and September 21, 2018 and September 21, 2018 to be included in this scoping review.

## 3.6 Step 4: Charting the Data

I entered extracted data into a form developed in the database programme *Microsoft Excel, Version 16.0, Windows*. The charting process was guided by the review question listed above in section 3.1.2. More specifically, data pertaining to the following headings were extracted for each included study: (1) authors; (2) year of online publication; (3) country or region that the study is from; (4) type of health professional that the BLP is focused on; (5) description or purpose of the BLP (i.e. to teach a 3$^{rd}$ year physiology course); (6) how the study labeled their BLP (e.g. blended learning or flipped learning); (7) components of usability that were being evaluated; (8) was the overall concept of usability evaluated or discussed; (9) methods of evaluation and analysis; and lastly, (10) if a questionnaire/survey was used, was it reliability tested, standardized, and or validated? The appropriateness of the charting form was discussed through consultation with my Thesis Advisory Committee (ChR, TC, TS) and academic liaison librarians.

Upon completion of charting, a fellow MSc learner (MHT) was recruited to validate the chart. MHT was asked specifically to confirm that data was accurately charted, add additional information or detail if needed, wherever possible, and ensure that language was uniform. MHT did this for all included studies.

## 3.7 Step 5: Collating, Summarizing, and Reporting the Results

### *3.7.1 Descriptive Quantitative Analysis*

Charted data were synthesized through tabulation. Descriptive quantitative analysis was then conducted on the extracted data present in the charted data (i.e. number of studies by year, how labels ascribed to BLPs have changed over the years, country of publication and the way these countries label their BLPs, and how the frequency of usability component evaluation has changed over the years) to describe the nature and distributions of the trends found through tabulation. These results are displayed in graphs and tables found within the results section below.

### *3.7.2 Qualitative Content Analysis*

To understand which components of usability were evaluated by each study, qualitative content analysis was utilized as described by Hsieh & Shannon, 2005 [73]. Specifically, directed content analysis was applied. The definitions of effectiveness, efficiency, and satisfaction as discussed by ISO 9241:11-2018(en) were used as the basis for developing initial codes (refer to Table 1 below). The findings of the qualitative content analysis were embedded into Table 2 found in the results section. Findings were validated by MHT.

**Table 1**: ISO Definitions of Usability Components and Examples of Coding

| Usability Component | Definition by ISO 9241-11:2018 | Example of what the component looks like in a BLP evaluation | Example of what would be coded under each component |
|---|---|---|---|
| Effectiveness | Accuracy and completeness with which users achieve specified goals | Measure of knowledge increase (i.e. grade change through pre-post test) | "This BLP assisted in my understanding of the content at hand" |
| Efficiency | Resources used in relation to the results achieved | According to ISO, “typical resources include time, human effort, costs and materials” when discussing efficiency | "I watched all the modules from beginning to end" |
| Satisfaction | Extent to which the user's physical, cognitive and emotional responses that result from the use of a system, product or service meet the user’s needs and expectations | Statement of enjoyment/disappointment with aspects of the BLP | "I was satisfied with this program" |

### *3.7.3 Qualitative Thematic Analysis*

Following the descriptive quantitative analysis, an inductive semantic qualitative thematic analysis was then conducted as discussed by Braun & Clarke, 2006, which was also followed by a deductive qualitative analysis using the ISO definition of usability [74].

All included studies were imported into QSR’s NVivo 12. Data was inductively coded by MHT and me independently. Both of us first utilized a substantive coding method where data (content throughout the articles that met eligibility criteria) was fractured based on new ideas, rather than focusing specifically on sentences or paragraphs. The rationale for this is that ideas related to usability components could be expressed by authors through both single sentences and

multiple paragraphs. Building on this thought, data was considered from all sections of the articles. As there is no explicit structure to discussing BLP evaluations, interesting thoughts related to usability components could appear throughout the article. MHT and I met at the 10% mark and the 36% mark to ensure consistency in findings. After completing initial coding, we met again to discuss our findings. If issues between codes were found, they were resolved through discussion. Following this, a brief discussion took place between us on potential themes that exist in the data. I then reviewed and organized all nodes into a hierarchical structure where sub-categories were grouped into overall categories, which grouped into sub-themes, which ultimately grouped into overarching themes. I then met with my primary supervisor (ChR) to discuss my interpretation of the data, and the overall themes that I decided to keep. Themes were further discussed and finalized in a committee meeting with my entire Thesis Advisory Committee (ChR, TC, and TS).

# 4. Results

## 4.1 Studies that met Eligibility Criteria & Charted Data

The search strategy was adapted to and implemented in Scopus and ERIC (EBSCO), yielding a total of 8626 titles after the removal of duplicates. Title and abstract screening, conducted via the use of a questionnaire guide, assisted in removing 8118 studies, leaving 508 studies to be full-text reviewed. After full-text screening, 53 studies were found to match the eligibility criteria, leaving 455 studies to be excluded. Of these 455 studies, 360 studies were primarily excluded because they did not meet the requirements of what a BLP should include – i.e. no learning management system was utilized, or the program did not apply asynchronous online learning material for 30 to 79% of the educational program. Seventy-one studies were excluded because they did not pertain to health professionals as defined by this study (e.g. the study was evaluating a BLP for veterinary learners). The remaining 24 studies were excluded because they did not conduct an empirical evaluation (n = 16) or did not evaluate for the concept of usability or one of its components (n = 8). Figure 1 provides a diagrammatic summary of the included and excluded studies. Table 2 presents a detailed summary of all extracted data from all studies that met eligibility criteria.

**Figure 1:** PRISMA flow diagram of included/excluded studies

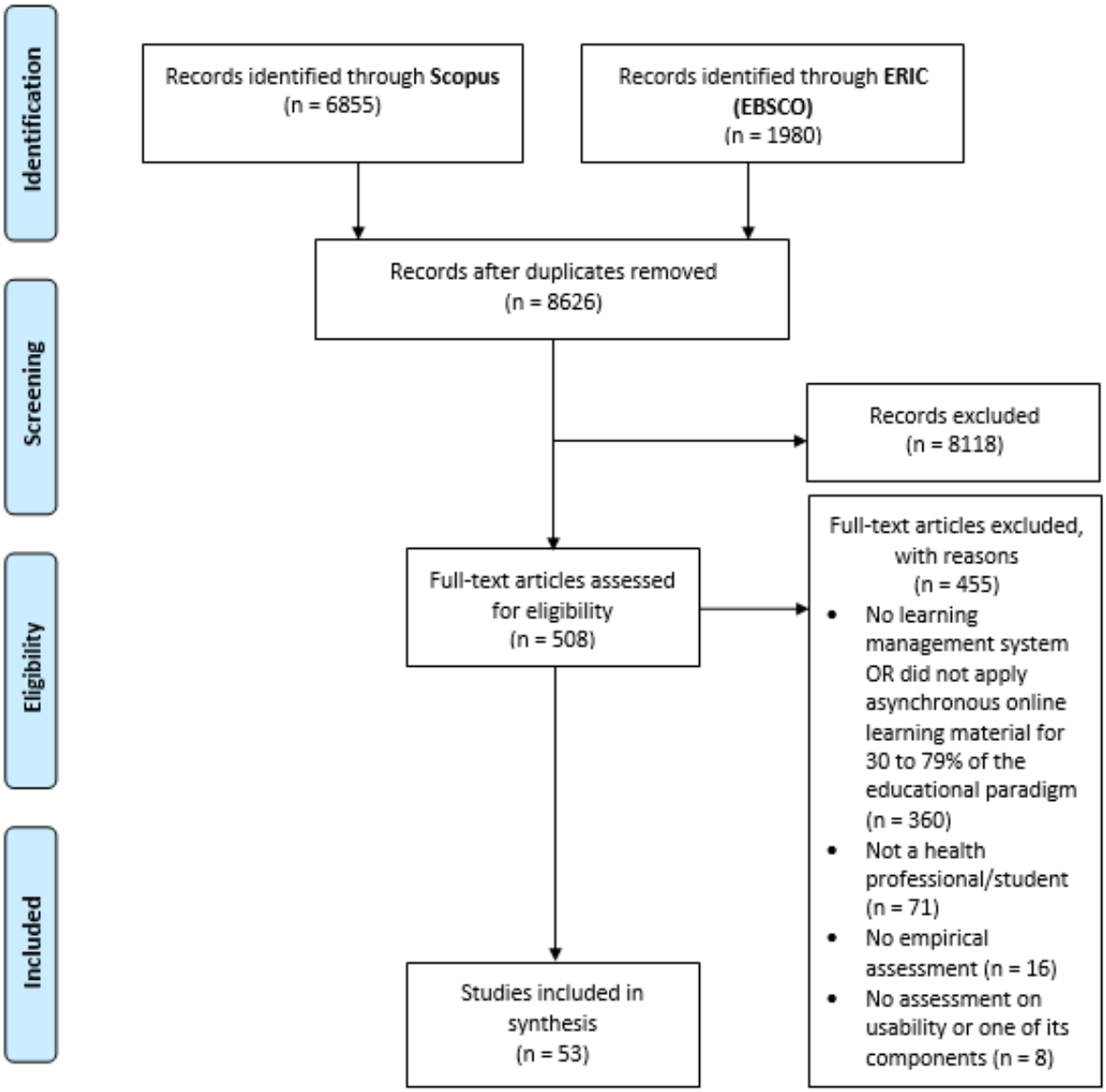


**Table 2:** Synthesized Table of Extracted Data for Studies that Met Eligibility Criteria

| Authors/In-text Citation | Type of Health Professional/ Learner | Description/Purpose of the BLP | How Does the Study Label their BLP? | Country/ Region of Study | Usability Component Being Evaluated | Was the overall concept of "Usability" Explicitly Evaluated/ Discussed? | Method of Evaluation and Analysis | Was Questionnaire/ Survey Reliability Tested, Standardized, and/or Validated if utilized? |
|---|---|---|---|---|---|---|---|---|
| **Alonso-Sardón et al., 2015** | Medical students (6th Year); Practicing family doctors and interns | To improve professional competence in the certification of causes of death in the Spanish National Health System | Blended Learning | Spain | Effectiveness; Satisfaction | No | "Quasi-experimental" pre- and post-survey. | Not Indicated |
| **Bohaty, Redfrod, & Gadbury-Amyot, 2016** | Dental students (2nd Year) | To teach pediatric dentistry | Flipped Classroom | United States | Effectiveness; Satisfaction; Efficiency | No | Pre- and post-survey (8 items), which includes an additional 12 questions administered at the end of the program only. Surveys included Likert scales (1-5). Thematic analysis conducted on open-ended questions. | Not Indicated |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Bonnes et al., 2017** | Medical residents (internal medicine) | To teach students in an internal medicine residency program | Flipped Classroom | United States | Effectiveness; Satisfaction; Efficiency | No | Kirkpatrick's model of evaluation level 1 and 2. Pre-post survey with Likert scale (1-5); comparison between students enrolled in the flipped classroom curriculum vs. the traditional course. | Study functioned to validate their survey |
| **Bösner, Pickert, & Stibane, 2015** | Medical students (4th and 5th Year) | To teach students differential diagnosis | Inverted Classroom | Germany | Effectiveness; Satisfaction; Efficiency | No | Pre- and post-survey (standardized questionnaire form University of Marburg) with Likert scales (1-5); focus group. | Standardized |
| **Bossaer et al., 2016** | Pharmacy students (3rd Year) | To teach pharmacotherapy oncology | Flipped Classroom | United States | Effectiveness; Satisfaction | No | ANCOVA analysis on examination scores using previous academic performance variables (i.e. undergraduate GPA) as covariates. Summative | Not Indicated |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | teaching evaluation (two-item questionnaire). | |
| **Boysen-Osborn et al., 2016** | Medical students (1st Year) | To teach advanced cardiac life support | Flipped Classroom | United States | Effectiveness; Efficiency | No | Comparison of three written evaluations (multiple choice questions) between students taking part in BLP vs traditional program. Ungraded 10-question quizzes to gauge student compliance with podcast viewing. | - |
| **Choi et al., 2015** | Nursing students (1st Year) | To teach a course called "Human Beings and Health" | Flipped Learning | South Korea | Effectiveness; Satisfaction; Efficiency | No | Flipped Course Evaluation Questionnaire; open-ended questions; focus groups. Conventional content analysis. | Not Indicated |
| **Crawford et al., 2013** | Practicing nurses | To teach a course called "Patient Navigation in Oncology Nursing" | Blended Learning | Canada | Effectiveness | No | Questionnaire (adapted from standardized questionnaire), including a Likert scale (1-5) and | Standardized |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | additional open-ended questions. | |
| **Domínguez et al., 2017** | Medical students (4th Year) | To teach a course on management of trauma patients | Inverted Classroom | Colombia | Effectiveness; Satisfaction; Efficiency | No | Pre- and post-test; generic institutional questionnaire; and evaluation by Flipped Classroom Perception Instrument (FCPI). | Not Indicated |
| **Duque et al., 2013** | Medical students (3rd and 4th Year) | To teach students in a geriatric medicine rotation | Blended Learning | Australia | Effectiveness; Satisfaction; Efficiency | No | Pre- and post-knowledge assessment instrument. | Not Indicated |
| **Eachempati, Kumar, & Sumanth, 2016** | Dental students (3rd and 4th Year) | To teach topics from dental pharmacology related to oral lesions and orofacial pain | Blended Learning | Malaysia | Effectiveness; Satisfaction; Efficiency | No | Qualitative thematic analysis of student reflections. | - |
| **Edwards, Kitzmiller, & Breckenridge-Sproat, 2012** | Emergency department staff members (RNs, nursing assistants, and unit coordinators) | To provide staff with health information technology training | Blended Learning | United States | Effectiveness; Satisfaction; Efficiency | No | Qualitative analysis of 13-question survey, including Likert scale (1-5). Responses summarized into "satisfaction score", plus additional thematic analysis | Indicates that the survey used was pre-existent, but does not use the word validated |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Elebiary & Al Mahmoud, 2013** | Nursing students (undergraduates – year not explicitly stated) | To teach a course about information technology for nurses | Blended Learning | Saudi Arabia | Effectiveness; Satisfaction; Efficiency | No | 6-tool descriptive research design, including comparison of student grades (enrolled vs. not enrolled in BLP); Student Satisfaction Survey (Likert scale 1-5), and teacher/course evaluations. | Indicates that the survey used is a modified version of the Students' Evaluation of Educational Quality Survey |
| **Forrest, 2010** | Practicing midwives | To provide midwives with increased training and education in perinatal mental health education | Blended Learning | United Kingdom (Scotland) | Effectiveness; Satisfaction; Efficiency | No | A modified online Objective Structured Clinical Examination (OSCE); Evaluation of portfolios of reflective accounts. | - |
| **Furnes, Kvaal, & Høye, 2018** | Nursing students (3rd Year) | To assist learners in strengthening their communication skills in mental health nursing | Blended Learning | Norway | Effectiveness; Satisfaction; Efficiency | No | Exploratory design; questionnaire with open-ended questions (Likert scale 1-5). Content analysis. | Determined face validity of the questionnaire through discussion with a reference group |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Gilboy, Heinerichs, & Pazzaglia, 2015** | Nutrition/dietetics students (undergraduates – year not explicitly stated) | To teach the courses "Professional Skills in Dietetics" and "Community Nutrition" | Flipped Classroom | United States | Effectiveness; Satisfaction; Efficiency | No | Survey with Likert scale (1-5) | Reliability Tested (assessed the Cronbach alpha of the survey prior to utilizing it) |
| **Goodie et al., 2011** | Medical students (3rd Year - family medicine clerkship) | To teach behavior change counselling | Blended Learning | United States | Effectiveness | No | Attitude and Knowledge assessment; 12-item pre- and post-class assessment. Additional 5 questions only at conclusion. | Not clearly indicated - mentions that items were derived from questions developed by Martino et al., 2007 |
| **Gopalan & Klann, 2017** | Pharmacy students (1st Year) | To teach advanced physiology | Flipped Teaching | United States | Effectiveness | No | Comparison of exam grades between different cohorts (flipped vs. non-flipped). | - |
| **Gostelow et al., 2018** | Medical students (4th Year) | To teach social determinants of health | Flipped Learning | United Kingdom | Effectiveness; Satisfaction; Efficiency | No | First level of Kirkpatrick's evaluation model: questionnaire (Likert scale 1-4); semi-structured group interview; thematic analysis. | Not Indicated |
| **Green & Whitburn, 2016** | Physiotherapy students (2nd Year) | To teach gross anatomy | Blended Learning | Australia | Effectiveness; Satisfaction; Efficiency | No | Retrospective cohort study of student grades and student | Not Indicated |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | feedback ("Likert-style questions"). Thematic and content analysis. | |
| **Howlett et al., 2011** | Medical student (5th Year) | To teach radiology content | Blended Learning | United Kingdom | Effectiveness; Satisfaction; Efficiency | No | Questionnaire. | Not Indicated |
| **Kangwantas et al., 2017** | Pharmacy students (2nd Year) | To teach the principles of nutrition for diabetes mellitus | Flipped Classroom | Thailand | Effectiveness; Satisfaction; Efficiency | No | Test scores compared between different cohorts (flipped vs. non-flipped); student feedback via 15-item survey (Likert scale 1-5); plus open-ended feedback from two peer instructors (not affiliated with course development or instruction). | Not Indicated |
| **Koo et al., 2016** | Pharmacy students (2nd Year) | To teach a pharmacotherapy course | Flipped Classroom | United States | Effectiveness; Satisfaction; Efficiency | No | Pre- and post test scores compared between different cohorts (flipped vs. un-flipped); pre- and post-course survey (designed to | Indicates that survey questions were adapted from a validated survey instrument of student attitudes toward televised courses |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | assess levels of Bloom's Taxonomy of Learning with Likert scale 1-5). Content analysis. | |
| **Kühl et al., 2017** | Medical students (1st Year) | To teach biochemistry | Inverted Classroom | Germany | Effectiveness; Satisfaction; Efficiency | No | Exam marks compared between different cohorts; course evaluation; questionnaire. Indicates that qualitative data was collected but did not reference the type of qualitative analysis that was conducted in the study. | Not Indicated (but does mention that the questionnaire was derived from a previous study by Rindermann et al., 2001) |
| **Langenau, Lee, & Fults, 2017** | Osteopathic medicine students (3rd Year - pediatric clerkship) | To teach students in a pediatric rotation | Blended Learning | United States | Effectiveness; Satisfaction; Efficiency | No | Osteopathic Medical Achievement Test (120 items) scores and final course grades; preceptor evaluations (18 items, Likert scale 1-10) were compared between the standard | Not Indicated |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | learning and blended learning groups; post-course survey. Identifies themes but does not discuss the type of qualitative analysis (i.e. thematic, content, etc.) or provide a reference to the approach that was used to derive these findings. | |
| **Leikola et al., 2009** | Practicing pharmacists | To assist pharmacy practitioners in acquiring competency in and accreditation for conducting collaborative comprehensive medication reviews (CMRs) | Does not refer to BL or any of its synonyms | Finland | Effectiveness; Satisfaction; Efficiency | No | Evaluation of participants' learning through learner diaries; written assignments and portfolio. Post-intervention survey (Likert scale 1-5) | Not Indicated, but survey routinely used by University of Kuopio, Centre for Training and Development |
| **Lorimer & Hilliard, 2009** | 2nd Year students in a Diagnostic Radiology and Imaging BSc Honours Program | To teach a course in relation to radiology and imaging | Blended Learning | United Kingdom | Effectiveness; Satisfaction; Efficiency | No | Two questionnaires, one for students and one for staff. Identifies themes but does not discuss the type | Not Indicated |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | of qualitative analysis (i.e. thematic, content, etc.) or provide a reference to the approach that was used to derive these findings. | |
| **Makhdoom et al., 2013** | Medical students (4th Year - family medicine course) | To teach a family medicine course | Blended Learning | Saudi Arabia | Effectiveness; Satisfaction; Efficiency | No | Dundee 'ready educational environment measure' (50-items, Likert scale 1-4); the 'objective structured clinical examination'; written examination with multiple-choice questions; analysis of case scenarios - comparison between intervention and non-intervention groups. | Validated |
| **Mary, Julie, & Jennifer, 2014** | Midwifery students (1st Year) | To teach a course on Research, Evidence and Clinical Practice | Blended Learning | Australia | Effectiveness; Satisfaction; Efficiency | No | University-based course evaluations (Likert scale 1-5) | Not Indicated |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Matsuda, Azaiza, & Salani, 2017** | Nursing students (accelerated undergraduates – year not explicitly stated) | To teach a course on evidence-based nursing practice. | Flipped Classroom | United States | Effectiveness; Satisfaction; Efficiency | No | 2 Surveys (one after pre-class module, one at end of semester) with Likert scale (1-5); plus qualitative questions. Conventional content analysis. | Not Indicated |
| **McLaughlin et al, 2014** | Pharmacy students (1st Year) | To teach the course Basic Pharmaceutics II (PHCY 411) | Flipped Classroom | United States | Effectiveness; Satisfaction; Efficiency | No | Pre- and post-course surveys. | Not Indicated |
| **McLaughlin et al., 2015** | Pharmacy students (2nd Year) | To teach Venous thromboembolism (VTE) to students enrolled in a pharmacotherapy course. | Blended Learning | United States | Effectiveness; Satisfaction; Efficiency | No | Comparison of engagement and performance based on online module access; pre-and post-test; response to in-class Automated Response System; exam performance; survey. | Not Indicated |
| **Milic et al., 2016** | Medical students (3rd Year) | To teach introductory medical statistics | Blended Learning | Serbia | Effectiveness; Efficiency | No | Comparison of grades (20 multiple choice test, plus final knowledge test) between students taking part in a blended | - |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | program and a traditional program. | |
| **Mohanna, Waters, & Deighan, 2008** | General Practice Trainers | To teach a "Modular Trainers Course" which provided instruction on General Practice Specialty Registrars | Blended Learning | United Kingdom | Satisfaction; Efficiency | No | Participant feedback (Likert scale 1-4). Identifies themes but does not discuss the type of qualitative analysis (i.e. thematic, content, etc.) or provide a reference to the approach that was used to derive these findings. | Not Indicated |
| **Moraros et al., 2015** | Master of Public Health students | To teach an introductory graduate course on epidemiology | Flipped Classroom | Canada | Effectiveness; Satisfaction; Efficiency | No | Surveys containing both Likert scale (1-5) and open-ended questions, which were administered at 3 time points. Additional Learner Evaluation of Educational Quality (SEEQ) Survey. | Standardized |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Morton & Colbert-Getz, 2017** | Medical students (1st Year) | To teach a required integrated basic-science course called Foundations of Medicine | Flipped Classroom | United States | Effectiveness; Satisfaction; Efficiency | No | Comparison of final exam marks between students that took part in the FC vs those in the LC; learner evaluations. Evaluations derived from Bloom's taxonomy. | Not Indicated |
| **Ocak & Topal, 2015** | Medical students (1st and 2nd Year) | To teach an anatomy course | Blended Learning | Turkey | Effectiveness; Satisfaction; Efficiency | No | Focus Groups with purposive sample of students with high, medium and low academic scores. Content analysis. | - |
| **O'Connor et al., 2016** | Medical students (3rd and 4th Year - radiology clerkship or elective) | To teach neuroimaging content | Flipped Learning | United States | Effectiveness; Satisfaction; Efficiency | No | 19-item electronic survey; shortened version of the class-related emotions section of the Achievement Emotions Questionnaire; pre- and post-test. | Validated |
| **Oh et al., 2017** | Nursing students (undergraduates– year not | To teach a flipped learning nursing informatics course | Flipped Learning | South Korea | Effectiveness; Satisfaction; Efficiency | No | 3 levels of Kirkpatrick's evaluation model: 10-item | Not Indicated |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | explicitly stated) | | | | | | questionnaire; course outcomes achievement (multiple choice test, essay, checklist); follow-up survey. | |
| **Ohtake et al., 2018** | Health professional students (dental medicine; dietetics; medicine; occupational therapy; pharmacy; physical therapy; social work; speech language pathology) | To teach an interdisciplinary evidence-based practice course | Flipped Classroom | United States | Effectiveness; Satisfaction | No | Module quizzes; Readiness for Interprofessional Learning Scale (RIPLS), plus survey with Likert scale (1-5). | Validated |
| **Page et al., 2017** | Allied health students (nursing; health science; podiatry; occupational therapist; physiotherapist; paramedicine; speech | To teach a first-year, first semester, physiology course | Blended Learning | Australia | Effectiveness; Satisfaction; Efficiency | No | Student grades; cross-sectional survey (Likert scale 1-5). Thematic analysis. | Not Indicated |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | pathology; exercise physiology; oral health) | | | | | | | |
| **Peisachovich et al., 2016** | Nursing students (1st Year) | To teach a first-year course on health assessment | Flipped Classroom | Canada | Effectiveness; Satisfaction; Efficiency | No | Student grades; feedback (comparison between different cohort of students - intervention vs no intervention). | Not Indicated |
| **Phillips, Schumacher, & Arif, 2016** | Pharmacy students (1st, 2nd, and 3rd Year) | To teach three courses: (1) small ambulatory care; (2) cardiovascular pharmacotherapeutics; and (3) evidence-based medicine | Blended Learning | United States | Effectiveness; Satisfaction; Efficiency | No | Exam grades; student survey; and additional open-ended questions asked to faculty. Thematic analysis on open ended questions. | Not Indicated |
| **Popovic et al., 2018** | Medical students; dental students; pharmacy students | To teach physiology | Blended Learning | Montenegro | Effectiveness; Satisfaction; Efficiency | No | Comparison of grades on assessments between intervention and non-intervention group; survey (Likert scale 1-5); use of online material. | Not Indicated |
| **Rose et al., 2016** | Emergency medicine residents (Post | To teach a course on pediatric emergency medicine | Flipped Classroom | United States | Effectiveness; Satisfaction; Efficiency | No | Pre- and Post-test; survey. Zaption analytics to | Not Indicated |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Graduate Year 3) | | | | | | determine levels of interaction with online content. | |
| **Rose et al., 2018** | Emergency medicine residents | To teach a course on pediatric emergency medicine | Flipped Classroom | United States | Effectiveness; Satisfaction; Efficiency | No | Pre- and post-test; survey. Zaption analytics to determine levels of interaction with online content. | Not Indicated |
| **Salajegheh et al., 2016** | 1st and 2nd Year medical students | To teach radiology interpretation skills | Blended Learning | Australia | Effectiveness; Satisfaction; Efficiency | No | Pre- and post-test; survey. Identifies themes but does not discuss the type of qualitative analysis (i.e. thematic, content, etc.) or provide a reference to the approach that was used to derive these findings. | Not Indicated |
| **Stewart et al., 2013** | Undergraduate medical students (year not explicitly stated - pediatric clerkship) | To improve newborn examination skills/neonatology | Blended Learning | Australia | Effectiveness; Satisfaction; Efficiency | No | Performance of newborn examination on standardised assessment compared between blended learning and | Not Indicated |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | control group; questionnaire. | |
| **Tang et al., 2017** | Medical students (4th Year ophthalmology clerkship) | To teach an ophthalmology clerkship | Flipped Classroom | China | Effectiveness; Satisfaction; Efficiency | No | A questionnaire modified from Paul Ramsden's Course Experience Questionnaire and Biggs' Study Process questionnaire with verified reliability and validity. | Adapted from a questionnaire that was previously reliability tested and validated |
| **Ward et al., 2011** | Practicing nurses | To teach a course on occupational health nursing | Blended Learning | United States | Effectiveness | No | Survey. | Not Indicated |
| **Wong et al., 2014** | 1st Year pharmacy students | To teach 3 classes on cardiac arrhythmias | Flipped Teaching | United States | Effectiveness; Satisfaction; Efficiency | No | Exam scores; 15-item survey (Likert scale 1-4); student feedback (comparison of intervention and non-intervention groups). | Not Indicated |
| **Ye & Smith, 2015** | Nursing students (senior level students – year not explicitly stated) | To teach a course on sleep education | Does not refer to BL or any of its synonyms | United States | Effectiveness; Satisfaction; Efficiency | No | Pre- and post-quiz; student feedback (Likert scale, 1-10). | Not Indicated |
| **Zomorodi et al., 2017** | Graduate health professional | To teach a course on population health and clinical emersion | Blended Learning | United States | Effectiveness; Satisfaction | No | Pre- and post-assessment; reflection | Validated |

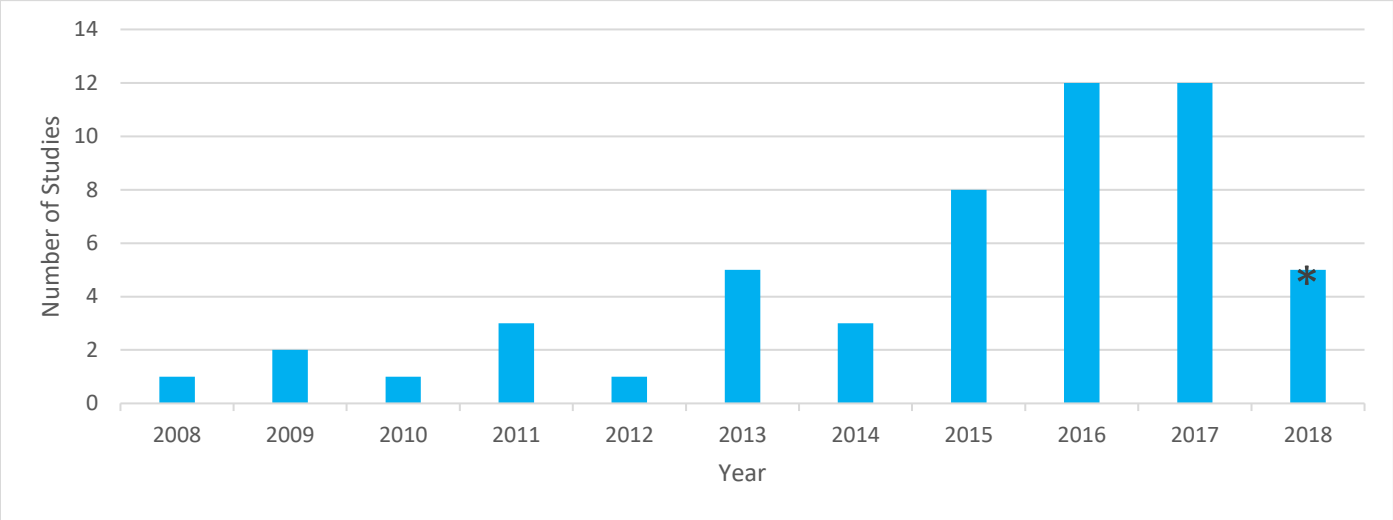
| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | students [medicine; nursing (clinical nurse leader); pharmacy; public health; and Master of Social Work] | | | | | | paper; assessment for Interprofessional Team Communication scale (AITCS) with Likert scale (1-5); course evaluations; benchmark reported through electronic medical record. | |

*Note that only studies that were published and indexed prior to September 21st, 2018 were included in this study.

## 4.2 Quantitative Analysis of Extracted Data

The oldest study that was included in this review was published in 2008. An overall increasing trend was identified in relation to the number of studies published each subsequent year. Please see Figure 2 below.

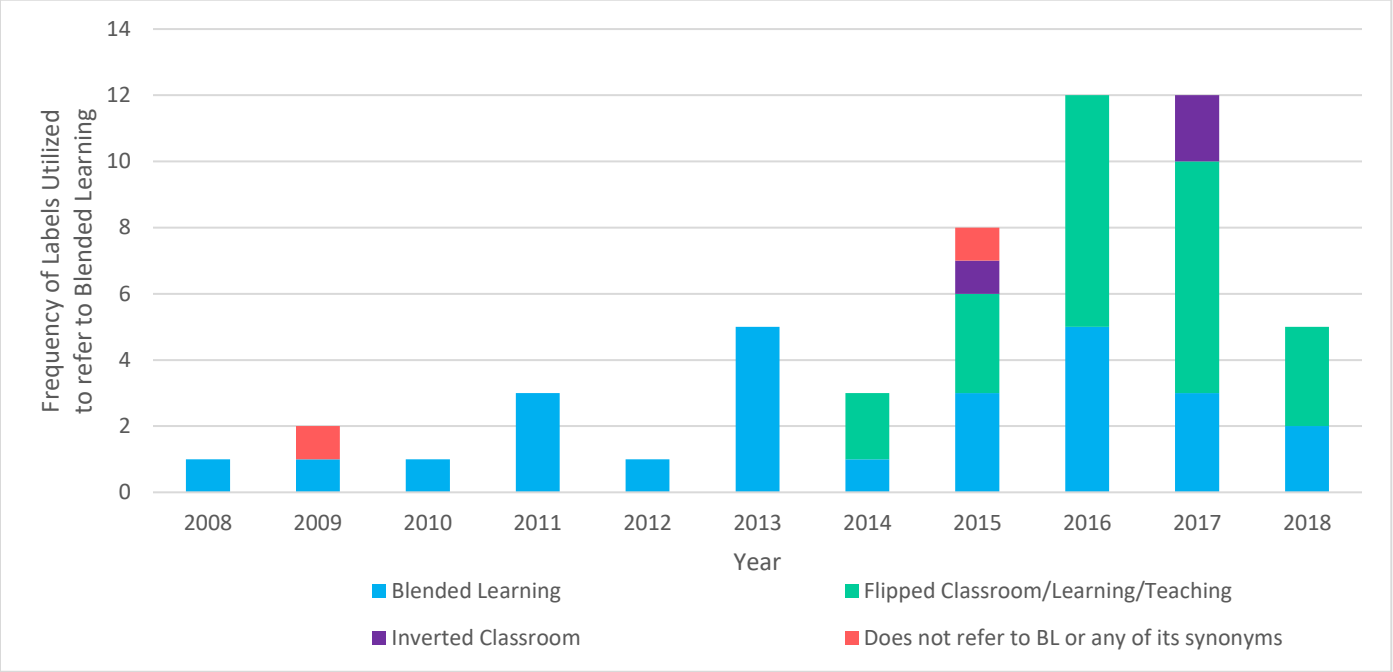
**Figure 2:** The number of studies that met eligibility criteria by year of publication




*Note that only studies that were published and indexed prior to September 21st, 2018 were included in this study.

Included studies looked at BLPs for a plethora of different populations in the field of health professions education including: medical learners, nursing learners, pharmacy learners, practicing midwives, and healthcare staff. Studies also came from around the world: North America, South America, Europe, Asia, and Oceania. When looking at the trend for naming BLPs, the label Blended Learning was exclusively used between 2008 and 2013. From 2014, studies began to primarily utilize the *Flipped Learning/Classroom/Teaching* model, rather than the general *Blended Learning* model. Please see Figure 3 below.

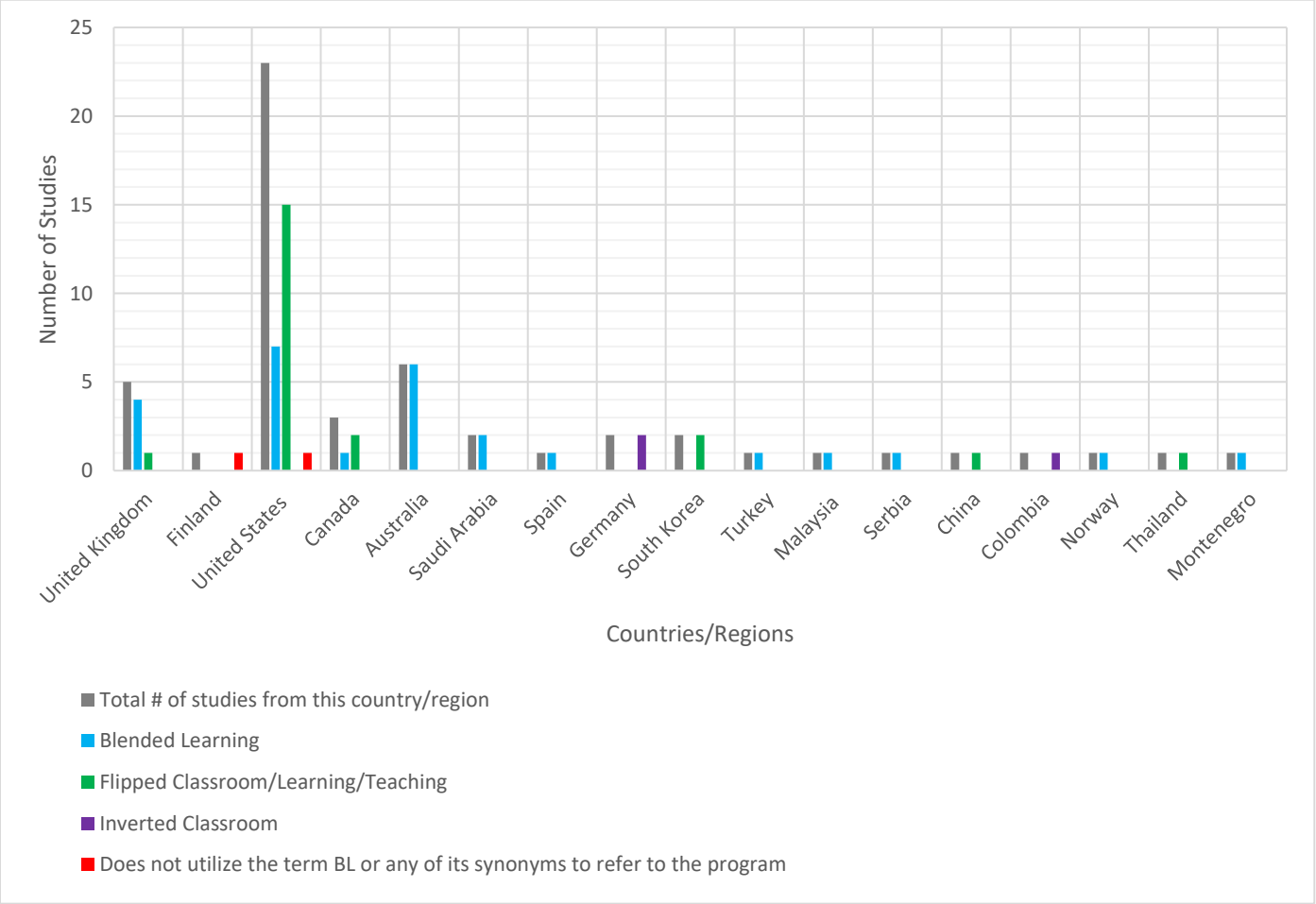
**Figure 3:** The frequency of labels used to refer to "blended learning" by year




*Note that only studies that were published and indexed prior to September 21st, 2018 were included in this study.

Most included studies came from the United States, followed by Australia, and then the United Kingdom. Most studies from the United States also employed the use of a *Flipped Learning/Classroom/Teaching* model. Australia, the United Kingdom, primarily use the term "Blended Learning." Only Germany and Columbia used the term "Inverted Learning." Please see Figure 4 below for an overview of the total number of studies that were included from each country/region, and how these studies were named in their respective countries/regions.

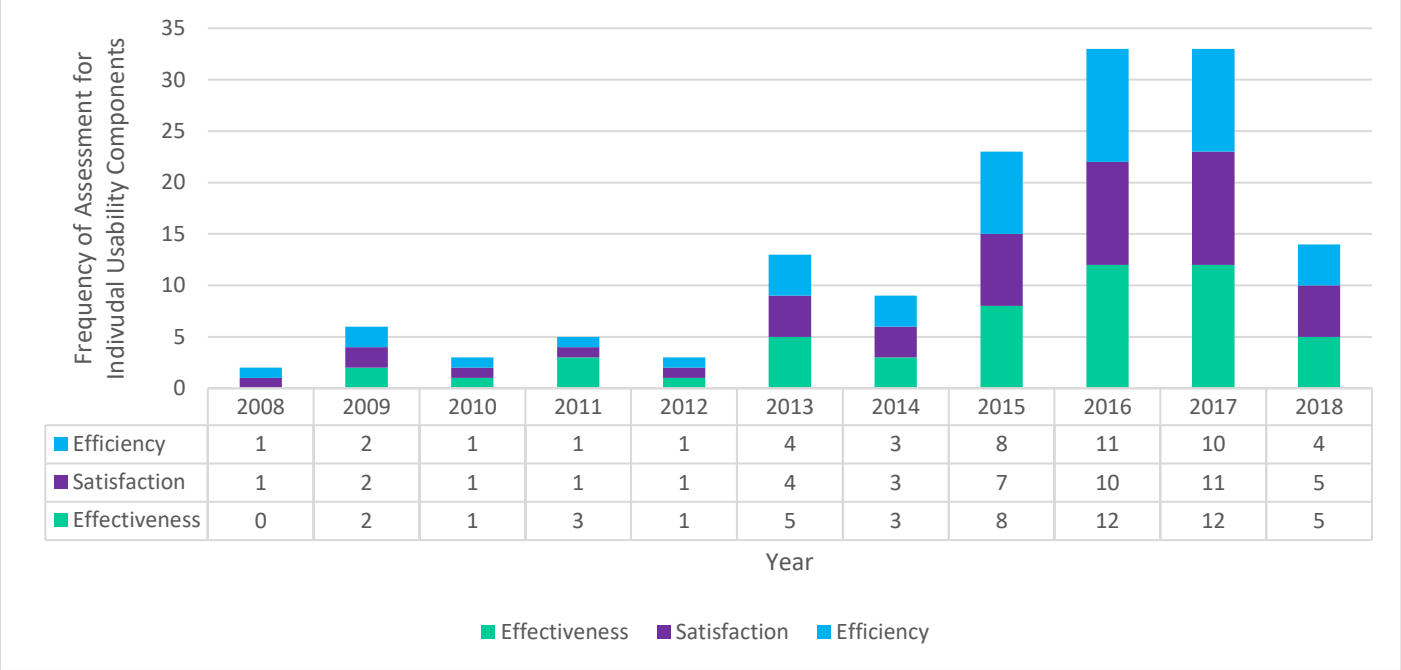
**Figure 4:** Total number of studies that met eligibility criteria from each country/region and the labels used to refer to blended learning by each respective country/region




*Note that only studies that were published and indexed prior to September 21st, 2018 were included in this study.

Importantly, no study that met eligibility discussed or evaluated the overall concept of "usability." When visualizing the frequency of evaluations conducted on individual usability components (effectiveness, efficiency, and satisfaction) identified through qualitative content analysis, no significant trend was identified. Please see Figure 5 below.

**Figure 5:** Frequency of usability components being evaluated by year

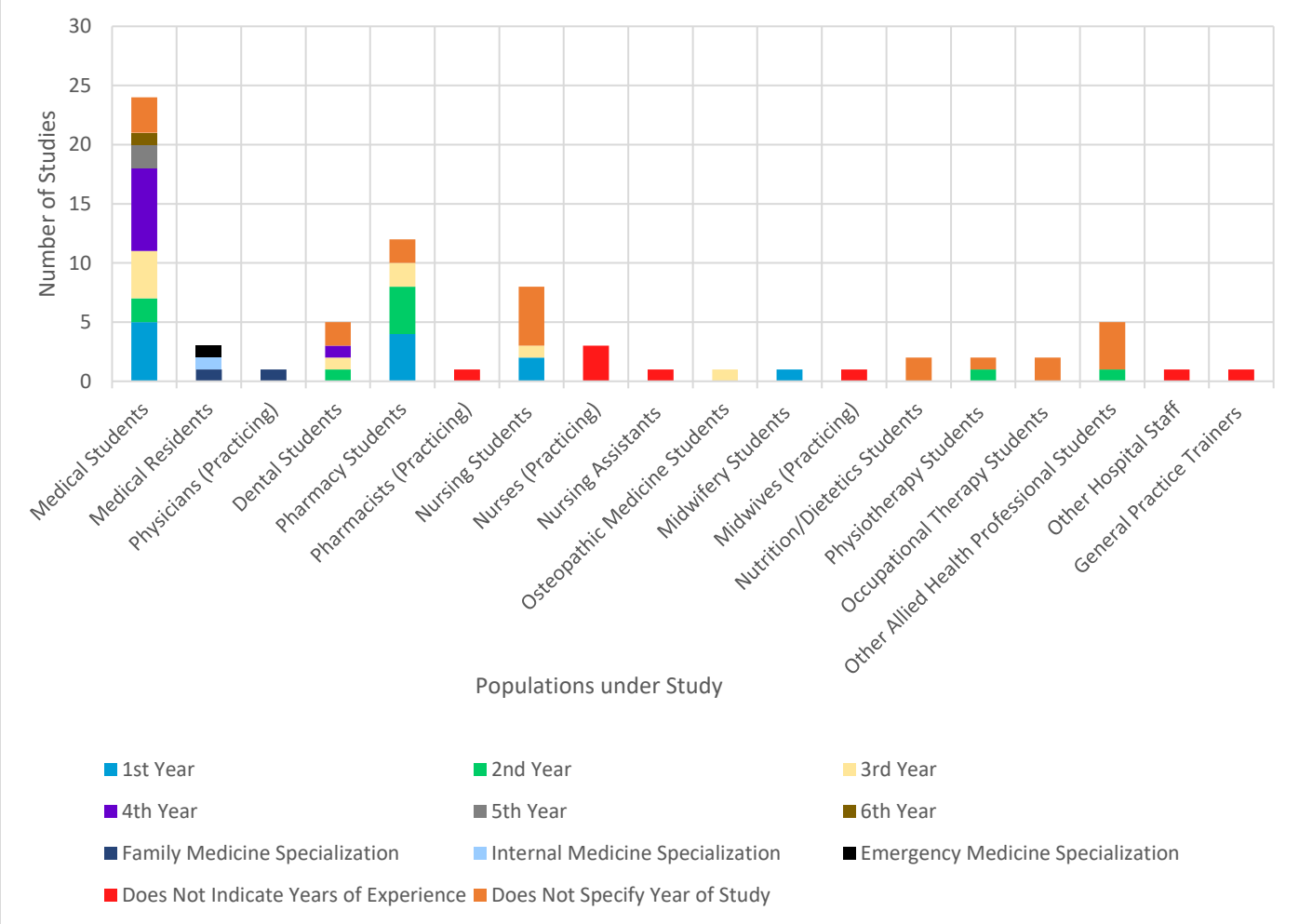



| | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Efficiency | 1 | 2 | 1 | 1 | 1 | 4 | 3 | 8 | 11 | 10 | 4 |
| Satisfaction | 1 | 2 | 1 | 1 | 1 | 4 | 3 | 7 | 10 | 11 | 5 |
| Effectiveness | 0 | 2 | 1 | 3 | 1 | 5 | 3 | 8 | 12 | 12 | 5 |

*Note that only studies that were published and indexed prior to September 21st, 2018 were included in this study.

When visualizing the frequency of studies by their specific population cohorts, it can be noted that most of the studies that met eligibility criteria focus on medical learners (24/53), followed by pharmacy learners (12/53), and thirdly, nursing learners (8/53). Please see Figure 6 below.

**Figure 6:** Visualization of the frequency of studies by their specific population cohorts




*Note that only studies that were published and indexed prior to September 21st, 2018 were included in this study.

Forty-seven out of the 53 studies (88.7%) utilized some sort of survey/questionnaire or course evaluation/feedback tool. Most of these instruments measured learner perceptions of the BLP they took part in (i.e. how satisfied were learners with their experience participating in the BLP). At times, studies also included evaluations on instructor/faculty perceptions (i.e. how enjoyable or useful they perceived the program to be). Twenty-five out of the 47 studies (53.2%) that used instrument, made explicit use of the Likert scale.

Thirty-one out of 53 studies (58.5%) explicitly indicated that a change in learner attitudes, knowledge, skills, and/or overall learning was evaluated. These evaluations were often done through a pre-post quiz/test design. Many studies utilized a combination of methods to measure the overall effects of BLPs.

Out of the 47 studies that utilized some sort of survey/questionnaire, or course evaluation/feedback tool, only 17 (36.2%) studies discussed to some extent if their instrument was reliability tested, standardized, or validated. More specifically, 4 studies (8.5%) clearly indicated that they utilized a validated instrument. Three studies (6.4%) indicated that their tool was standardized. One study (2.1%) conducted a reliability evaluation (evaluation of Cronbach Alpha) prior to the application of their instrument. One study (2.1%) indicated that the survey that was utilized was routinely used by the university in which the evaluation took place but did not mention if appropriate research was conducted to standardize or validate the survey. One study (2.1%) determined the face validity of the questionnaire through discussion with a reference group. One study (2.1%) indicated that it was conducted to validate their instrument. Three studies (6.4%) indicated that their instrument was adapted/modified from one that was previously validated and/or reliability tested. Three studies (6.4%) indicated that the survey that was used was either pre-existent or adapted from a previously developed instrument but did not mention if the instrument

was previously reliability tested, standardized, or validated. Among these 17 studies that indicated or discussed validation, reliability testing, or standardization, tools utilized in these studies were not identified as being specifically developed to evaluate BLPs. Rather, these tools were developed to measure concepts such as "communication" or "learning" in general, and not specifically within the context of a BLP.

Three out of the 53 studies (5.6%) applied focus groups as a method of BLP evaluation. One out of 53 studies (1.9%) utilized semi-structured group interviews as its method of BLP evaluation. Interestingly, 7 studies (13.2%) applied content analysis and 6 studies (11.3%) applied thematic analysis to data collected through open-ended questions, learner feedback, or learner reflections. Five studies (9.4%) identified themes from their data or mentioned that qualitative data was collected in their study but did not reference the type of qualitative analysis technique that was used.

Most studies that met eligibility criteria did not follow or refer to any formal evaluation framework when evaluating their BLP. Only 3 studies (5.6%) [84, 100, 120] referenced the 4-level Kirkpatrick model of evaluation [75, 76]. Out of these 3 studies, one completed only the 1$^{st}$ level of evaluation, another completed the 1$^{st}$ and 2$^{nd}$ level of evaluation, and the last study completed the 1$^{st}$ through 3$^{rd}$ levels of evaluation.

## 4.3 Qualitative Thematic Analysis

All retained materials for this scoping review were coded independently by two reviewers, MHT and AA. The reviewers met three times during the coding process – at the 10%, 36% and 100% marks. Each meeting confirmed that both reviewers had analyzed the materials similarly, in that the codes were alike in nature and frequency. In the final meeting, potential themes that could be interpreted from our readings of the data were discussed. Following this, the initial codes were restructured into three hierarchical themes based on Braun & Clarke's framework for conducting thematic analysis [74]. Themes were further developed and refined through discussions with my thesis committee members and particularly with my primary supervisor ChR. Following the initial inductive semantic thematic analysis of the studies retained in this scoping review, a deductive analysis was carried out to assist in further analyzing and bring structure to the data.

## Part 1: Inductive Analysis

### Theme 1: Avoiding the 'Usability' Label and Using Undefined Related Terms

As noted in the quantitative synthesis developed above, one of the more striking themes emerging from the qualitative analysis is that none of the works retained for this review employed the term "usability" when reporting BLP evaluations. Having in mind the definition of usability adopted in this study, I found that several other terms and concepts that can be related to usability in BLPs have been used across studies; for instance: "appropriate" [94, 104, 108]; "clear" or "clarity" [101-105, 110, 111, 118, 120, 128, 129, 132]; "easy" or "ease-of-use" [88, 102, 118, 125, 129, 132]; "efficacy" [105, 110, 114, 119]; "favourable" [88]; "flexibility" [88, 91, 93-102, 108, 109, 111, 114, 118, 122, 124, 125, 127, 128, 130]; "helpful" [101, 105, 130, 132]; "informative" [96]; "useful" or "usefulness" [85, 86, 88, 89, 91, 100, 105, 106, 110, 115, 118, 119]; "utility" [111]; and "worthwhile" [111]. What is more, whereas each of these 'alternative' terms have a specific connotation, interpretations of these words could be quite ambiguous, as illustrated by the following excerpt:

> "Using a 4-point Likert scale (strongly agree, agree, disagree, strongly disagree), students were asked to rate the *helpfulness* of the pre-class lecture recording, clarity of learning objectives, and *overall effectiveness* of the flipped method" [italics are mine] [132].

Here, the word *helpfulness* seems to be initially related to the idea of assisting learners in their understanding. However, without a clear definition or reference provided by the authors, the word *helpfulness* could have many other subjective interpretations, depending on various factors, i.e. learners' mood when completing the evaluation. Furthermore, when conducting evaluations without clear definitions, learners may also use synonyms to assist in their understanding of

specific terms. In this case, *helpfulness* could be regarded as a synonym for *usefulness* or *beneficial*, just to name a few. Once again, both synonyms have specific connotations and can be interpreted differently when clear definitions or frameworks are not present.

*Ease-of-use* is often considered to be a misrepresentation or a highly simplified version of the concept of usability [21]. Interestingly, as mentioned above, *ease-of-use* was one of the terms/concepts that was consistently discussed across multiple studies. When discussing this concept, the authors applied it exclusively to the technological components of their BLPs. Still, some level of ambiguity towards this concept exists as studies either fail to define the term or they appear to ascribe different connotations to it. Refer to the following two excerpts:

> "[Online case discussions] were introduced to provide the students with the opportunity to discuss case scenarios with a moderator–the forum needed to be *easy to use*, available to all students, to support images and video and to be secure from non-module user access. The solution was provided by an open-source social networking plugin to the [Virtual Learning Environment] …" [italics are mine] [102].

And:

> "*Easily used* media platforms may also influence student satisfaction. In the feedback, many students stated preferring voice-over PowerPoint slides to the WebEx video recordings because they found the former easier to navigate (i.e. find/repeat specific content), and the PowerPoint slides received fewer complaints regarding technical issues than the WebEx video recordings received" [italics are mine] [132].

In the first excerpt, *ease-of-use* appears to be discussed as a separate concept from that of availability and access. Whereas in the second excerpt, *ease-of-use* is linked to the concept of

navigation and technical difficulties. Again, without the provision of an explicit definition or a framework that guides the use of terminology, ambiguity in the interpretations of terms and concepts by authors is shown to result.

Another example that is worth highlighting is the confusion around the term *useful* or the concept of *usefulness*. Unlike the term *ease-of-use* which was applied only to technological aspects of BLPs, the term *useful was* applied across various aspects of BLPs. The way in which the word *useful* and the overall concept of *usefulness* was applied and interpreted seemed to differ between learners and authors across studies. Refer to the following two excerpts:

> "The students preferred learning the course online since the crowded setting decreased the laboratory's effectiveness and access to the resources out of the class was *useful* for their learning. 'It was like a course for us. It is like we are taking a class for 1 or 2 more hours at home. From this point, it was very *useful*'" [italics are mine] [118].

And:

> "After completion of their ICA in geriatric medicine, 88% of the students agreed that WebCT was a *useful* tool for this rotation. When the students were asked about their perceptions of the use of a paper-based portfolio, 68% agreed that they felt *comfortable using it* whereas 16% somewhat disagreed with this statement" [italics are mine] [91].

In the first excerpt, the authors discuss the concept of *usefulness* as related to that of *accessibility* based on what a learner has reported. When reading the learner's perception however, the connotation of the term *useful* becomes relatively unclear. Their sentence seems to be referring to both the *accessibility* and the increased amount of *time* the learner can spend learning the material at home. In the second excerpt, the term *useful* is however linked to the idea of comfort

in use. This could be considered in line with the term *ease-of-use* as discussed earlier. Therefore, once more, terms and concepts, when left undefined or referenced allow for ambiguity of interpretations to arise, and in this case, not just between authors of studies, but also between the interpretations of authors and participants of the same study.

## Theme 2: Confusing Conceptualization of the Components of Usability (i.e. Effectiveness, Efficiency, & Satisfaction)

Whereas the term usability was not used across the retained studies for this review, terms referring to the specific components of usability do appear consistently. The conceptualization of these terms, however, differs significantly across studies. This theme functions to capture and analyze the ambiguity surrounding the way in which these terms were discussed across the studies that were retained for this review.

### *Effectiveness*

*Effectiveness* was explicitly used in most of the studies that were analyzed for this review. Although widely applied, the interpretation of this word was seldom unanimous. No study provided a framework to define this term, but often associated the term with unique ideas or concepts specific to each study. For instance, refer to the following two excerpts:

> "Our results are in agreement with those of other studies on the *effectiveness* of e-learning as part of blended learning, which showed that *students' engagement* was increased, and their *perception* of the educational environment was improved. The only domain that was affected negatively by blended learning in our study was the social perception. Thus, although the use of technology in teaching is *effective* and is *perceived* as such, it requires a cultural change in learning practice that might not be easy for everyone" [italics are mine] [109].

And:

> "When the findings on the *effective learning* of [Blended Learning Environment] were analyzed, students stated that the images made learning *long-lasting*; made learning *easy*

for the students; and *helped* the students get prepared before the class … The students preferred learning the course online since the crowded setting decreased the laboratory's *effectiveness* and access to the resources out of the class was *useful* for their learning …" [italics are mine] [118].

In the first excerpt, *effectiveness* is discussed as related to the concepts of *learner's engagement* and *perception*. Whereas in the second excerpt, the concept of *effectiveness* is related to several different concepts including: *permanence of learning*, *ease-of-learning*, and *assistance with pre-class preparation*. The second excerpt goes on to discuss the effectiveness of laboratory settings, however, the interpretation of what the authors of the second excerpt are saying here is quite confusing even to us as there seems to be a lack of punctuation. Regardless, there seems to be a link between *effectiveness* and *the number of learners taking part in a laboratory activity*. Ultimately, these two excerpts function nicely in illustrating how authors across studies interpret the term *effectiveness* differently.

### *Efficiency*

Although studies were often discussing the concept of *efficiency* as can be seen through the results of the qualitative content analysis shown above (refer to Table 2), studies did not often apply the label of *efficiency*. In the few instances that the term was explicitly applied, the connotation that authors apply to it differ. For instance, refer to the following two excerpts:

"Although a *well crafted* and *captivating* lecture presentation seems like an *efficient* way for an instructor to cover course content, converging evidence implies that listening to a classroom lecture is not an effective way to promote deep and lasting student learning" [italics are mine] [99].

And:

> "Students frequently claim that they prefer podcasts to real-time instruction because they can both *speed up* the podcast, running at 1.5 or 2X speed, as well as *review* portions of podcasts that they need to see again. They view this as more *efficient*" [italics are mine] [87].

When analyzing these two excerpts, the connotation that authors ascribed to the concept of *efficiency* can be seen to differ drastically. Whereas the first study relates the idea of "well-crafted" and "captivating" lectures to efficiency, the second study discusses *efficiency* as related to the concept of *time* (i.e. speed of learning) and the ability to review parts of lectures.

However, it should be noted that in the retained studies for this review, what was found to be much more common was that studies often discussed concepts associated with *efficiency* (i.e. time, human effort, cost) without adding any labels to their descriptions. Refer to the following excerpt for an illustration:

> "The *time* required by students to complete the asynchronous online lectures was less than the time allocated for the lectures for three of the five lectures … The magnitude of the difference ranged from *12-22 minutes* under the allocated time, which means that students spent *24-28% less* than the *allocated time* for these lectures. For two of the lectures in the elective course, the students spent *more time* (range 3-5 minutes, representing 4-6% of allocated contact hours), but this difference was not significant. *Student-reported time* to complete the online lectures differed from the time measured by Articulate Online for all of the online lectures in all of the courses. In all instances, students estimated that it took them *longer* than it actually did to complete the online lecture. Students overestimated the *time* to completion by an average of *24.8 minutes per hour lecture*" [124].

These types of statements were common across the retained studies for this review. Therefore, with regards to *efficiency*, I have identified that there is not only a poor application of this label, but a significant amount of ambiguity regarding its connotation.

### *Satisfaction*

*Satisfaction* was discussed across most studies included in the included studies; the word "satisfaction" was nonetheless not always applied. Instead, the words "positive" and "negative" were often applied to indicate learner perceptions. The ambiguity in relation to this concept arises through two specific issues: (1) each study applied different connotations to the concept of *satisfaction* (i.e. different terms are related and ascribed to satisfaction across studies); and (2) studies differ regarding what their focus of evaluation was: either on the *satisfaction* of specific components of the BLP or the entire program in general. The following three excerpts properly illustrate this result:

> "*Satisfaction* has been widely used as one of the important parameters to evaluate learning *effectiveness* in academic institution. *Higher student satisfaction* is the *results of good learning*. In addition, the present study found significant differences in course *satisfaction* between the blended learning sections and the classical section, with blended students reporting a *higher level of class satisfaction*. The blended learning design focused on active learning in the classroom portion of the course; the students might have rated *higher satisfaction due to* the *enjoyment* of the blended design" [italics are mine] [94].

> And:

> "The *positive student evaluations and outcomes* of this course demonstrate that it is possible to *effectively engage* students in research courses when the right ingredients are

combined. The tangible and lucid link to practice throughout this course is clearly a key element in its success…" [italics are mine] [110]

And finally:

"A *poor attitude* toward online learning is not a reflection of *dissatisfaction* with the subject. Students have generally perceived the online quizzes as a *positive learning experience*. Survey results regarding online learning quizzes are consistent across all years of the survey." [italics are mine] [122].

As can be seen from these excerpts, the first two studies discuss *satisfaction* as a concept that is used to measure *effectiveness*, whereas the last study does not link *satisfaction* and *effectiveness* in the same way. Rather, *satisfaction* is discussed in this last study in relation to the concepts of *attitudes* and *experiences*. Furthermore, each excerpt indicates that studies either use different labels to orient the focus of evaluations when discussing *satisfaction*: the first discussed learner *satisfaction* regarding the BLP; the second discussed learner *satisfaction* regarding the course (which essentially refers to the BLP); and the last study discussed *satisfaction* regarding online learning and specifically the online quizzes.

### Theme 3: Lack of Consensual Approach to Evaluation

The qualitative analysis of articles selected for this scoping review has evidenced that conceptual and definitional ambiguity concerning usability and its primary components has been accompanied by an absence of a consensual approach to evaluating this phenomenon in the context of BLPs.

Studies did not explicitly discuss the methodological approach they undertook when completing their study. Furthermore, each study adopted a unique set of methods to evaluate these concepts (i.e. each study used a unique questionnaire/survey or evaluation/feedback tool). Each study also had a different focus of evaluation, where some studies put a greater emphasis on evaluating the e-learning component, and others focused more heavily on the in-person component, and others attempted to evaluate the BLP as a whole (i.e. rather than asking questions specific to each component of the BLP, questions were made more general – e.g. "did you enjoy your experience with the BLP?").

Many studies often attempted to complement subjective measures (i.e. learner perceptions found through questionnaires) with objective measures (i.e. changes in grades through pre-post tests). In this case, it seemed necessary for data collection methods to be unique because each BLP covered different content and so it would make sense for pre-post tests to be BLP-specific. Nevertheless, not every author decided it was necessary to measure their BLPs subjectively and objectively. For instance, some studies only looked at perceptions through questionnaires, whereas others only completed analysis on learner grades.

Tools utilized by authors differed in their items (i.e. the questions they asked), and in the way the items were evaluated. For instance, most studies used Likert Scales for their items, but the scales were either 1-4, 1-5, or 1-10 depending on the study. Some authors included open-ended

responses in their tools. These works varied in their analysis of these open-ended responses. In these cases, some studies were found to use thematic analysis, some used content analysis, and others discussed themes or findings for analyzing these open-ended responses without explicitly indicating which type of qualitative analysis method was utilized.

Many studies did not explicitly discuss if their tools were validated, reliability tested, and/or standardized. The tools that were utilized in each study differed in the effects that they were measuring. For example, in one study, scholars utilized a shortened version of a validated questionnaire called the *Achievement Emotions Questionnaire*. This was used to measure "achievement emotions related to the course" [119]. The items of this tool seemed relevant to some extent in measuring aspects of learner satisfaction and perceived effectiveness with the program in general. Another validated tool that was used in an included study by scholars was the *Evaluation for Interprofessional Team Communication Scale* [134]. This tool was used as a measure of team collaboration, which can be considered a potential measure of effectiveness.

As discussed in the quantitative analysis, three studies referenced the 4-level Kirkpatrick model of evaluation [84, 100, 120]. This model has been largely used for evaluating educational programs. However, when observing how this model was utilized in each of the three studies retained for this scoping review, results were disparate. Each of the three studies conducted different levels of evaluation. Although evaluations of *satisfaction* can be seen in each study, different tools were used to measure *satisfaction*. When analyzing the third study [120] that conducted evaluations on levels 1 through 3 of the Kirkpatrick model, measures of *satisfaction* and measures that can be considered as corresponding to *effectiveness* (based on ISO definitions) can be made.

Interestingly, authors indicate that reflection had taken place regarding how they planned to conduct their evaluations or why certain evaluation methods were or were not utilized. Refer to the following excerpt:

> "We are aware that the presence of a control group improves the scientific rigor of the study, but it was not used because the classic formative intervention has proven to be effective, and it would therefore be unethical to deprive a group of learners of Medicine and/or doctors of this model" [82].

In this regard, our analysis has helped identify seven "take home" ideas steaming from researchers' reflective stance regarding what BLP evaluations should consist of:

(1) Evaluations should address the in-class and online learning aspects separately;

(2) Evaluations should also consider the BLP as a whole;

(3) Control groups could be beneficial when comparing results within a study;

(4) A need exists to measure long-term changes in learners;

(5) Utilizing an evidence-based approach to evaluation is critical;

(6) Evaluations should consider both objective measures such as grades, and subjective measures such as perceptions on satisfaction;

(7) Multiple strategies should be used to comprehensively evaluated the effects of BLPs.

It should be noted that although these seven ideas were consistently discussed across studies, perceptions on these ideas or the ways in which they were evaluated differed significantly across studies, further indicating the lack of consensus regarding evaluation approaches for BLPs.

## Part 2: Deductive Analysis

### ISO Framework

Terms used in all studies that were retained for this scoping review seemed quite ambiguous, and the evaluations that were conducted appeared significantly disparate. I thus decided to apply a deductive approach to the thematic analysis after completing the inductive analysis, where the ISO framework for usability [21] guided further understanding of the data. Through this further analysis it can be noted that an implicit reference to the concept of usability was made throughout the included studies, and particularly in the studies that evaluated for more than one usability component, yet again without naming them as such. The following four definitions outlined in the ISO were used to guide this deductive analysis [21]:

(1) "Usability is an outcome of use" and is defined as "the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use."

(2) "Effectiveness is the accuracy and completeness with which users achieve specified goals" and it "represents the extent to which actual outcomes match intended outcomes."

(3) "Efficiency is the resources used in relation to the results achieved" where "resources are considered expendable" and "resources include: time, human effort, money and materials."

(4) "Satisfaction is the extent to which the user's physical, cognitive and emotional responses that result from use of a system, product or service meet user's needs and expectations." It is important to note that "satisfaction influences user behaviours and accomplishments."

Note that effectiveness, efficiency, and satisfaction can be evaluated as observable outcomes (objective measurements) or through user's perceptions of the outcome (subjective measurements) [21]. Also note that systems, products or services can be: effective without being efficient or providing users satisfaction; or efficient without being effective or satisfying; or provide users with satisfaction without being effective or efficient [21].

## Usability as a Means for Increased Consistency in Labelling

With these definitions and notes in mind, it became relatively easy for me to replace the different labels that authors used across studies with the above terms. For instance, in the quote "students also perceived that the FLM is a useful and favourable method. However, findings from previous studies of students' satisfaction with the FLM were inconclusive" [88], the word *useful* could potentially be replaced by *effective* and the term *favourable* could be replaced by *satisfying*. By replacing these words with terms that are defined and guided by an established framework, increased comparability and generalizability of results can be achieved because it would allow for all stakeholders (i.e. learners that complete questionnaires, researchers that administer questionnaires, and individuals that read and learn form these studies) to be on the same page in terms of what the words they are using mean, and how these terms could be evaluated for.

It could also be noted that applying ISO definitions for usability concepts enables authors to become significantly more succinct and transparent throughout their studies. For instance, in the following excerpt – which was previously discussed in theme 1 – the word helpfulness was interpreted to be abstract and relatively ambiguous:

> "Using a 4-point Likert scale (strongly agree, agree, disagree, strongly disagree), students were asked to rate the helpfulness of the pre-class lecture recording, clarity of learning objectives, and overall effectiveness of the flipped method" [132].

When using the ISO framework, *helpfulness* could be interpreted as an overarching concept which combines perceived effectiveness, efficiency, and satisfaction. If the author was to utilize these three terms independently, greater clarity for learners completing the survey and for authors that attempt to compare and learn from this study's evaluation methodology can be achieved.

Continuing, the following excerpt indicates how the concept of usability can allow for increased consistency regarding the definitions that authors ascribe to their terms and thereby, what constitutes and evaluation for that term:

> "Our results are in agreement with those of other studies on the effectiveness of e-learning as part of blended learning, which showed that students' engagement was increased, and their perception of the education environment was improved. The only domain that was affected negatively by blended learning in our study was the social perception" [109].

When reading this excerpt, the authors indicate that effectiveness is determined by: (1) learners' engagement, a concept that seems to fit well under the usability component of efficiency as it related to the idea of human effort; and (2) their perceptions of the educational environment, a concept that seems to fit well under the usability component of satisfaction. This would mean that the author would evaluate for effectiveness as a measure of increased satisfaction. ISO indicates that effectiveness, efficiency, and satisfaction are independent, but related concepts [21]. This implies that perceived effectiveness may be affected by satisfaction and efficiency but is not entirely dependent on these concepts. Therefore, the authors in this case would need to further develop their analysis.

## Interpreting Usability from Results Sections – A Comprehensive Example

My deductive analysis of the works retained for this scoping review revealed that, taking into account terms and synonyms, all studies addressed at least one or more component of usability (namely effectiveness, efficiency, satisfaction), and at times, studies also addressed 'accessibility' and 'user experience', two concepts which ISO describes as being critical and closely related to usability [21]. The following excerpt illustrates how a typical study within the studies retained for this review discussed the results of their BLP evaluation, and more specifically, how the terms authors used could be interpreted as synonymous with the concept of usability and its related concepts.

> "The feature of the course that participants liked most was the eLearning modules. They found them very interactive, creative, easy to understand, and useful in addressing multiple learning styles. Participants appreciated the accessibility and self-paced nature of the eLearning modules. The participants also valued the peer-reviewed journal readings and reported that these readings complemented the material presented in the modules and reinforced current practice issues and evidence-based practice. Participants reported that the discussion forums, which were another interactive part of the course, allowed nurses an opportunity to share opinions, knowledge, and practice experiences" [89].

In this excerpt, several results are illustrated – each of which was identified through analysis as strongly associated with components of usability and its related concepts of user experience and accessibility. The usability concept of effectiveness can be interpreted from "easy to understand." Where perceptions regarding emotional, cognitive, and physical responses to the program are made (i.e. "participants liked …" and "participants valued"), reference to the usability concept of satisfaction can be interpreted. Where perceptions regarding the resources used in

relation to the results that are achieved are made (i.e. “readings complemented the material presented in the modules and reinforced…”), reference to the usability concept of efficiency can be interpreted. The concept of accessibility is discussed explicitly in the above excerpt. Where general learner perceptions regarding their interaction with the BLP are made (i.e. participants appreciated the … self-paced nature of the eLearning modules), the concept of user experience can be interpreted as being applied.

### Consistency in Application of Concepts to Evaluate Usability Components

When using the deductive approach to analyze the 53 studies that were retained for this review, 31 key concepts associated with each of the components of usability (i.e. effectiveness, efficiency, and satisfaction) and its related concepts (i.e. user experience and accessibility) were found to be consistently evaluated across studies. These key concepts include: change in knowledge, change in skills, perceptions, and cost-benefit analysis to name a few. Studies often did not label these concepts. Rather, many of them were interpreted through the items present in their evaluation instruments. To see a full list of these concepts and which usability components/related concepts they fit under, please see Appendix 4.

### Benefits of BLPs – Implicit References to Usability

Lastly, through deductive analysis, it could be noted that authors often utilize the components of usability and its related concepts of accessibility and user experience to discuss the benefits of BLPs in both the introduction and discussion/conclusion sections of their studies. For example, some studies discussed the ability of BLPs to decrease the overall cost to implement learning programs. This benefit can be interpreted as a reference to the efficiency of BLPs. Two other benefits that are consistently discussed across studies are that BLPs allow for greater

competency acquisition and that they facilitate meaningful learning. These benefits can be interpreted as a reference to the effectiveness of BLPs. Please see Appendix 4 for a list of what studies identified as the benefits of BLPs (coded as categories and sub-categories), and how some of these benefits were interpreted to fit within each of the usability components and its related concepts of user experience and accessibility (coded as sub-themes).

## 4.4 Summary of Findings from Qualitative Thematic Analysis

In summary, three themes were developed through thematic analysis: (1) Avoiding the 'Usability' Label and Using Undefined Related Terms; (2) Confusing Conceptualization of the Components of Usability (i.e. Effectiveness, Efficiency, & Satisfaction); and (3) Lack of Consensual Approach to Evaluation. Through these three themes, it can be noted that studies in the field of health professions education that conduct evaluations on BLPs utilize different ambiguous terms to discuss their evaluations. Where the same terms are used across studies, different interpretations are found to be ascribed to these words. All of this results in disparate evaluation methods across studies. However, when applying a deductive approach to qualitative analysis, the words and concepts authors used were found to fit well within the definitions for effectiveness, efficiency, and satisfaction, and thereby the framework for usability, as discussed by the ISO [21]. In doing so, greater clarity and consistency are brought to the terms and concepts, and this can assist in establishing consensus in BLP evaluation approaches across studies.

## 5. Discussion

The purpose of this scoping review was to develop a foundational understanding of how the concept of usability and its primary components have been defined and evaluated for in BLPs within the field of health professions education. The most important finding of this scoping review is that although the paramount importance of usability has been strongly established in relation to e-learning platform evaluation [19, 20, 26, 31-33, 77-79], no study was found to explicitly discuss or evaluate for usability in relation to BLPs in the field of health professions education. This indicates that a major lack in knowledge translation exists where the implementation of this imperative concept has yet to be achieved in the current field of inquiry. That said, this scoping review also allowed me to nuance this finding.

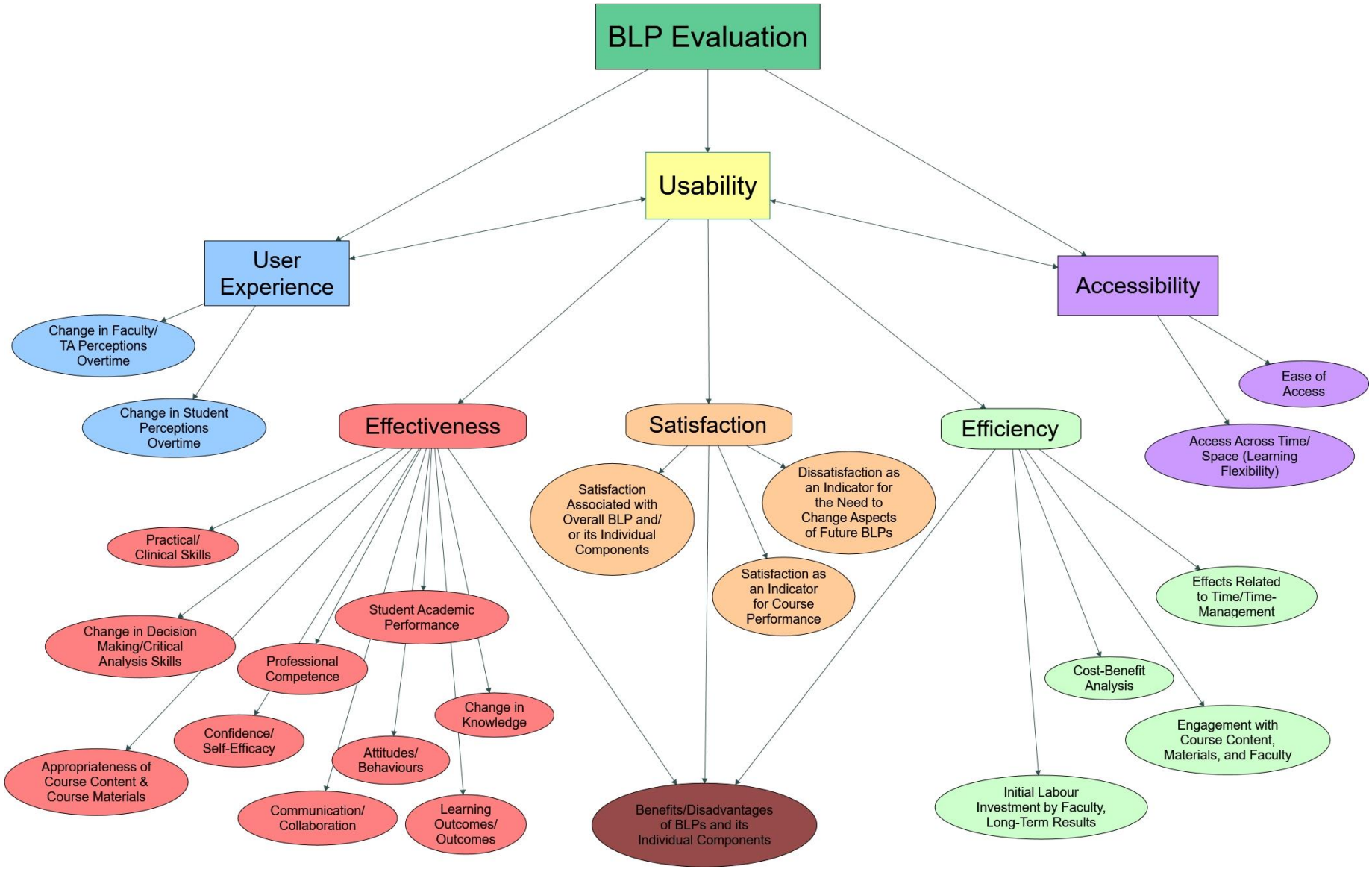
When focusing on BLPs, the results from the quantitative results of this study highlight that this pedagogy has been applied widely and diversely across the field of health professions education. Furthermore, most authors have measured perceived effects of the implementation of BLPs using various questionnaires/surveys or course evaluation/feedback tools, often without indicating if the evaluation instrument utilized was validated, standardized, or reliability-tested. These quantitative results suggest that no consensus amongst researchers exists regarding appropriate evaluation methods for BLPs.

Whereas the word usability was absent, it is important to stress that a second major result of this investigation, stemming from the results of the qualitative content analysis presented in Table 2, is that scholars of BLPs do evaluate for usability components *effectiveness, efficiency, and satisfaction* according to the definitions outlined by the ISO framework for usability [21]. This suggests that authors may have a general sense of what concepts are important to evaluate their BLPs for, but they lack the link to the overarching concept of usability.

Through the inductive thematic analysis conducted in this study, I was first able to elucidate the disparity and ambiguity in the terms that authors applied across the included studies. What is more, I identified the unique and abstract connotations that authors ascribed to usability components (i.e. effectiveness, efficiency, and satisfaction), as well as the large differences in the evaluation methods authors utilized across the studies. These findings not only confirm the lack of authors' consensus which was implied in the quantitative findings but also further the discussion developed through the prior qualitative content analysis, in that authors do indeed discuss usability components across studies without explicitly referring to the usability construct; and they do so in an abstract and unique manner from one another. Therefore, the third major result of this scoping review is that authors do not utilize a consistent framework or a pre-defined set of evaluation terminology. This may severely impact the comparability and generalizability of these studies.

The corollary of my analytic-synthetic work has been a deductive approach to perform analysis guided by ISO framework for usability and its components. This phase of the scoping review assisted in elucidating the implicit reference authors make to the concept of as identified in the previous inductive analysis portion of this review. In fact, in this body of knowledge, the deductive analysis assisted in identifying 31 concepts that were related to usability, its components (i.e. effectiveness, efficiency, and satisfaction), and importantly, its two related concepts (i.e. accessibility and user experience). To assist in analyzing the relationships between usability and these components and related concepts, I developed a concept map with the support of QSR's NVIVO 12. Through visualization, the 31 concepts identified in the deductive analysis were summed up into 22 critical concepts that were evaluated across the included studies. Please see Figure 7 below.

**Figure 7:** Concept Map Developed from the Deductive Findings of the Thematic Analysis



When reviewing the concept map, it is important to keep in mind that this figure is a depiction of the *critical concepts* that have been evaluated for in BLPs and not the actual *methods* that authors utilize to evaluate these concepts. For example, a study that evaluated a BLP for the concept of effectiveness could have done so through measuring "student academic performance" or through measuring a change in the "practical/clinical skills" of students. However, to measure these concepts, authors could employ various methods of their choosing.

This concept map clearly illustrates that, even without naming it explicitly, usability appears to be at the core of BLP evaluations. Even when authors evaluate for user experience or accessibility, they always relate their results back to the concept of usability through highlighting one of its key components (i.e. accessibility can be linked to the effectiveness of programs). Furthermore, this concept map also depicts that when studies discuss the benefits and disadvantages of BLPs or their individual components (i.e. online learning vs face-to-face learning), implicit reference to all three major usability components are often made.

The importance of this concept map is two-fold: (1) it assists in the theoretical development of usability and (2) the practical evaluation of this concept. Firstly, the development of this concept map functions to enrich the framework for usability outlined by ISO [21]. More specifically, through this map, a stronger understanding of how the components of usability (effectiveness, efficiency, and satisfaction) and its related concepts (user experience and accessibility) are associated with one another is achieved. The ISO framework briefly mentions the user experience and accessibility are concepts that do not fall under usability, but rather they are related to one another [21]. Furthermore, the framework mentions that when evaluations consist of all three of these concepts, they can be considered to fall under the over-arching concept of "human-centred quality and design" [21]. However, this concept map further elucidates the inter-relation between

these three concepts. In the include studies, when authors evaluated for user experience or accessibility, their results were always related back to the concept of usability, albeit without naming it as such. Secondly, the 22 critical concepts illustrated in this map assist in developing a clear understanding of how usability components and related concepts have been practically evaluated for in relation to BLPs within the current context. Therefore, this map will be of paramount importance for stakeholders that plan on conducting future evaluations of BLPs. Specifically, this concept map will assist in identifying the critical concepts that stakeholders would like to focus their BLP evaluations on (i.e. do students meet learning objectives and do they enjoy their time in the program) and how these concepts fit under the framework of usability (i.e. effectiveness and satisfaction).

Alongside the concept map, another finding of this thesis is the seven ideas regarding what BLP evaluations should consider – presented under theme 3 of the qualitative analysis portion of this review (refer to page 77). The importance of these seven "take home" ideas is that they provide a very clear guideline as to what a comprehensive BLP evaluation should include. Together, the concept map and these seven "take home" ideas will assist in the development and validation of an instrument and framework to evaluate usability in BLPs (pre and post-deployment of the BLP) within the field of health professions education.

In summary, BLPs have been employed widely and vastly across the field of health professions education. Although their evaluations focus on usability components and related concepts, albeit in a highly disparate manner, the term usability is never explicitly stated. However, the concept map and the seven "take home" ideas regarding BLP evaluations identified in this scoping review are instrumental in (1) enriching the framework for usability as outlined by the

ISO and (2) developing a tool to evaluate usability in the current context prior to BLP implementation as well as post-implementation.

## Limitations

Scoping reviews provide an opportunity to analyze a broad range of literature when the research question is exploratory in nature. Although scoping reviews assists in developing a foundational understanding of the topic of focus, the methodology does have some limitations. When screening titles, abstracts, and full-text articles, relevant studies can be overlooked due to human error. To counter this, it would be preferable to have two independent reviewers screen all titles, abstracts, and full text-articles. In this study, only one reviewer was used to screen all titles, abstracts, and full-text studies. However, co-authors and liaison librarians were consulted heavily throughout the screening process.

Furthermore, with the exponential increase in the number of studies published each year, literature reviews have the potential to lose significant power quickly. Rerunning the same search strategy four months after data collection was completed resulted in hundreds of new studies appearing in Scopus and ERIC (EBSCO). As studies that were only published and indexed in Scopus and ERIC (EBSCO) prior to September 21$^{st}$, 2018 were included in this review, relevant studies published and indexed after this date may have been missed. Moreover, studies were only retrieved from two databases. Although we believe this was sufficient, searching the grey literature could have been beneficial.

Lastly, quality appraisals of included articles are not typically conducted when completing a scoping review [64]. Therefore, this study does not discuss the robustness or generalizability of the findings in each included article. However, as the focus of this review was on developing an

understanding of the extent and nature of research activity and identifying any gaps in the literature, the lack of a quality appraisal was acceptable.

## Strengths

A major strength of this study is that it includes articles from all areas of health professions education research. This means that the search strategy utilized in this study attempted to retrieve articles on both students and professionals in the field of healthcare that received any educational training (i.e. a course in an undergraduate nursing program or a continuing professional development initiative) through a BLP. Figure 6 depicts the broad range of population cohorts that were considered in this study, whereas Figure 4 depicts the broad range of countries that the included studies came from. It is important to note that a large portion of the population falls under the more specific domain of medical education (i.e. medical students). Furthermore, population cohorts included in this study are associated with many levels of care including primary (i.e. family doctors and nurses) and secondary (i.e. occupational therapists, physiotherapists, and dentists). Therefore, the findings of this study have the potential to impact the educational training of a vast array of individuals within the field of healthcare.

That being said, it is important to highlight the clarity that the deductive analysis brought to the large ambiguity and disparity in the data. This strongly indicates the ability of the ISO framework for usability to bring consensus into the body of evaluation literature on BLPs within the field of health professions education. This in turn, will allow for the further strengthening of BLPs in this field and will potentially lead to better healthcare delivery globally.

## Contributions and Future Directives

This scoping review has identified the extent, range, and nature of BLP evaluations in the field of health professions education in relation to usability and its primary components. It has also functioned to illustrate the vast disparity and ambiguity around the terminology and evaluation methods that authors have utilized in their studies. However, this study also demonstrates the ability of the ISO framework for usability to bring clarity and consensus into the literature. Two major outcomes of this study are (1) seven "take home" ideas that discuss important factors to consider when conducting BLP evaluations and (2) a concept map that depicts the relationship between usability, its components (effectiveness, efficiency, and satisfaction), its related concepts (user experience and accessibility), and 22 critical concepts that were identified through deductive analysis of the literature. The seven "take home" ideas and the concept map will be used in conjunction to guide my next steps which are to develop and validate an instrument to evaluate usability in BLPs within the field of health professions education.

Along with all these contributions, we wish to highlight several concerns identified through this study which may lead to future areas of research, or at the very least, identify questions that may be important to consider in the future.

Firstly, the fact that many studies were excluded in the screening phase of this review due to not meeting the requirements of what a BLP should include (i.e. an LMS), prompts the question of how BLPs are being defined in the field of health professions education. As a reminder, this study only touched on how articles label their BLP (refer to Table 2). It would also be interesting to see if BLPs benefit through the application of a rigid definition as presented in this study. These questions may necessitate a systematic review to address them.

Secondly, the abstruse and varied application of terminology and dissimilar connotations that authors ascribed to terms suggests the need for either: 1) a guideline or lexicon of terms and their meanings to be developed for this context; or 2) the adoption of a developed framework such as the ISO framework for usability. Through these, future researchers would be able to ascertain clarity and consistency in terms of their use of terminology and concepts across studies.

Continuing, the disparate methods that authors applied when evaluating their BLPs suggests the need to conduct a systematic review to comprehensively search for any tools or frameworks that can evaluate the synchronous face-to-face components and asynchronous online learning components of BLPs in a comprehensive and time-efficient manner prior to their wide implementation in health professional education settings. Should no appropriate tool or framework be found, then a need to develop one exists. The concept map and the seven "take home" ideas of what a BLP evaluation must include, identified through this study, will be of paramount importance in this endeavour – one that I plan on taking on as my next steps as mentioned above.

Thirdly, as technology continues to advance, it would be interesting to see future studies discuss the application of the concept of usability on BLPs that incorporate online learning means for asynchronous and synchronous teaching and learning components (i.e. the use of videoconferencing technology to incorporate face-to-face instruction with asynchronous online learning methods).

## Contributions Specific to Family Medicine

Family medicine education research is a relatively underdeveloped area of inquiry within the field of health professions education [80, 81]. However, as a discipline, family medicine is a crucial medical speciality, integral to primary care delivery, which in turn is critical in ensuring

that health care systems function effectively and efficiently [81, 135, 136]. As health care systems around the world continue to evolve to better accommodate changing health care priorities in a rapidly developing technological era [80, 135], so to do our educational systems need to advance in order to ensure physician competency amidst these changes [80, 135]. This is imperative for family physicians in particular because scholars indicate that a primary healthcare revolution is impending in which family physicians will play a pivotal role [135]. With this in mind, educational programs using BL and directed towards family medicine faculty development have begun to emerge globally [137-140]. The knowledge developed through this scoping review will help ensure that programs such as these are evaluated appropriately, thereby advancing knowledge in family medicine, strengthening primary care delivery, and safeguarding the robustness of health care systems around the world.

## 6. Conclusion

To my knowledge, this is the first study that applies a rigorous strategy to search for evaluation studies in the field of health professions education, with the use of a strong definition for blended learning and usability. Several major findings are identified in this scoping review study. Quantitative descriptive analysis highlighted the extent, range, and nature of the included studies. Inductive thematic analysis of the included studies functioned to confirm the large disparity found in the quantitative analysis and to further shed light on the ambiguity regarding the conceptualization of the various terms used by authors across the included studies. Whereas deductive thematic analysis then assisted in developing an understanding of how the ISO framework of usability can bring consistency in labelling and interpretation across the studies. Alongside the importance of usability and the overall ISO framework for usability in the current context, a concept map and seven key ideas of what BLP evaluations should include were identified through this study – both of which will be of paramount importance for stakeholders that plan on conductive BLP evaluations in the future or developing a new framework for evaluating BLPs using the concept of usability. Thus, this study functions to provide a foundation of knowledge that can potentially strengthen BLPs in the field of health professions education, and possibly the wider field of education research. Notably, this study has the potential to strengthen health professions education across all levels of care (i.e. students and professionals affiliated with different levels of healthcare) in general, and family medicine education in particular, which in turn will function to potentially strengthen healthcare delivery.

## References


1. Garrison, D. R., & Kanuka, H. (2004). Blended learning: Uncovering its transformative potential in higher education. *The internet and higher education*, 7(2), 95-105.
2. Garrison, D. R., & Vaughan, N. D. (2008). *Blended learning in higher education: Framework, principles, and guidelines.* John Wiley & Sons.
3. Williams, C. (2002). Learning on-line: A review of recent literature in a rapidly expanding field. *Journal of Further and Higher Education*, 26(3), 263–272.
4. Vahed, M., Rakeeb Hossain, R. M., Yu, R., Javed, S., & Nguyen, T. (2017). The Future of Blended Learning: A Report for Students and Policy Makers, *Research Paper* 1617-005. Retrieved from: https://academicanswers.waldenu.edu/faq/73164
5. Graham, C. R. (2006). Blended learning systems. *The handbook of blended learning*, 3-21.
6. Graham, C. R., Woodfield, W., & Harrison, J. B. (2013). A framework for institutional adoption and implementation of blended learning in higher education. *Internet and higher education, 18(3)*, 4-14.
7. Frenk, J., Chen, L., Bhutta, Z. A., Cohen, J., Crisp, N., Evans, T., Kelley, P. (2010). Health professions for a new century: transforming education to strengthen health systems in an interdependent world. *The lancet, 376*(9756), 1923-1958.
8. Choules, A. (2007). The use of elearning in medical education: a review of the current situation. *Postgraduate Medical Journal*, 83(978), 212-216.
9. Bersin, J. (2004). *The blended learning book: Best practices, proven methodologies, and lessons learned*. John Wiley & Sons.

10. Ruiz, J. G., Mintzer, M. J., & Leipzig, R. M. (2006). The impact of e-learning in medical education. *Academic medicine*, 81(3), 207-212.
11. Masters, K., & Ellaway, R. (2008). e-Learning in medical education Guide 32 Part 2: Technology, management and design. *Medical teacher*, 30(5), 474-489.
12. Means, B., Toyama, Y., Murphy, R., & Baki, M. (2013). The effectiveness of online and blended learning: A meta-analysis of the empirical literature. *Teachers College Record*, 115.
13. Pereira, J. A., Pleguezuelos, E., Merí, A., Molina-Ros, A., Molina-Tomás, M. C., & Masdeu, C. (2007). Effectiveness of using blended learning strategies for teaching and learning human anatomy. *Medical education, 41*(2), 189-195.
14. Twigg, C. (2003). Improving Learning and Reducing Costs: Lessons Learned from Round I of the Pew Grant Program in Course Redesign. *Center for Academic Transformation, Renseelaer Polytechnic Institute. Toy: NY.*
15. Singh, H. (2003). Building effective blended learning programs. *Educational Technology-Saddle Brook Then Englewood Cliffs NJ-, 43*(6), 51-54.
16. Allen, I. E., Seaman, J., & Garrett, R. (2007). *Blending in: The extent and promise of blended education in the United States*. Sloan Consortium. PO Box 1238, Newburyport, MA 01950.
17. Stacey, E., & Gerbic, P. (2009). Introduction to blended learning practices. In *Effective blended learning practices: Evidence-based perspectives in ICT-facilitated education* (pp. 1-19). IGI Global.
18. Watson, J. (2008). Blended Learning: The Convergence of Online and Face-to-Face Education. Promising Practices in Online Learning. *North American Council for Online Learning*.
19. Sandars, J. (2010). The importance of usability testing to allow e-learning to reach its potential for medical education. *Education for Primary Care*, 21(1), 6-8.

20. Asarbakhsh, M., & Sandars, J. (2013). E-learning: the essential usability perspective. *The clinical teacher*, 10(1), 47-50.

21. International Organization for Standardization. (2018). ISO 9241-11:2018(en) Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts. Retrieved from https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en

22. Van Nuland, S. E., Eagleton, R., & Rogers, K. A. (2017). Educational software usability: Artifact or Design? *Anatomical sciences education*, 10(2), 190-199.

23. Bevan, N., Carter, J., & Harker, S. (2015). *ISO 9241-11 revised: What have we learnt about usability since 1998?* Paper presented at the International Conference on Human-Computer Interaction.

24. Frøkjær, E., Hertzum, M., & Hornbæk, K. (2000). *Measuring usability: are effectiveness, efficiency, and satisfaction really correlated?* Paper presented at the Proceedings of the SIGCHI conference on Human Factors in Computing Systems.

25. Bevan, N., Carter, J., Earthy, J., Geis, T., & Harker, S. (2016, July). New ISO standards for usability, usability reports and usability measures. In *International Conference on Human-Computer Interaction* (pp. 268-278). Springer, Cham.

26. Parlangeli, O., Marchigiani, E., & Bagnara, S. (1999). Multimedia systems in distance education: effects of usability on learning. *Interacting with computers, 12*(1), 37-49.

27. Hansen, G. K., Olsson, N., & Blakstad, S. H. (2010). *Usability evaluations–user experiences–usability evidence.* Paper presented at the CIB Proceedings: Publication.

28. Precel, K., Eshet-Alkalai, Y., & Alberton, Y. (2009). Pedagogical and Design Aspects of a Blended Learning Course. *International Review of Research in Open and Distance Learning, 10*(2).

29. Jeng, J. (2005). Usability assessment of academic digital libraries: effectiveness, efficiency, satisfaction, and learnability. *Libri, 55*(2-3), 96-121.

30. Seffah, A., Donyaee, M., Kline, R. B., & Padda, H. K. (2006). Usability measurement and metrics: A consolidated model. *Software Quality Journal, 14*(2), 159-178.

31. Ssemugabi, S., & De Villiers, M. R. (2010). Effectiveness of heuristic evaluation in usability evaluation of e-learning applications in higher education. *South African computer journal, 2010*(45), 26-39.

32. Sandars, J., & Lafferty, N. (2010). Twelve Tips on usability testing to develop effective e-learning in medical education. *Medical teacher, 32*(12), 956-960.

33. Zaharias, P., & Poylymenakou, A. (2009) Developing a usability evaluation method for e-learning applications: beyond functional usability. *International Journal of Human-Computer Interaction* 25: 75–98.

34. Stocking, J. E., & Mo, B. P. (1995). Current status of interactive multimedia education in medicine. *MD computing: computers in medical practice*, 12(5), 373-81.

35. Ellaway, R., & Masters, K. (2008). AMEE Guide 32: e-Learning in medical education Part 1: Learning, teaching and assessment. *Medical teacher*, 30(5), 455-473.

36. Hiemstra, R., & Sisco, B. (1990). Individualizing Instruction. Making Learning Personal, Empowering, and Successful. *Jossey-Bass, Inc., Publishers, 350 Sansome Street, San Francisco*, CA 94104.

37. Gotovsky, K., Lee, C., Bhimani, J., Ding, K., Vahed, M., Hossain, R., Min, R., Yu, R., Javed, S., & Nguyen, T. (2017). The future of blended learning: A report for students and policy makers. *Federation of Canadian Secondary Students*, 1-14.

38. Snart, J. A. (2010). Hybrid Learning: The Perils and Promise of Blending Online and Face-to-Face Instruction in Higher Education: The Perils and Promise of Blending Online and Face-to-Face Instruction in Higher Education. ABC-CLIO.
39. Dziuban, C., Moskal, P., & Hartman, J. (2005). Higher education, blended learning, and the generations: Knowledge is power: No more. *Elements of quality online education: Engaging communities. Needham, MA: Sloan Center for Online Education*, 88-89.
40. Strayer, J. F. (2012). How learning in an inverted classroom influences cooperation, innovation and task orientation. *Learning environments research*, *15*(2), 171-193.
41. Bergmann, J., & Sams, A. (2009). Remixing chemistry class: Two Colorado teachers make vodcasts of their lectures to free up class time for hands-on activities. *Learning & Leading with Technology*, *36*(4), 22-27.
42. Milman, N. B. (2012). The flipped classroom strategy: What is it and how can it best be used?. *Distance Learning*, *9*(3), 85.
43. Roehl, A., Reddy, S. L., & Shannon, G. J. (2013). The flipped classroom: An opportunity to engage millennial students through active learning strategies. *Journal of Family & Consumer Sciences*, *105*(2), 44-49.
44. McLaughlin, J. E., Roth, M. T., Glatt, D. M., Gharkholonarehe, N., Davidson, C. A., Griffin, L. M., . . . Mumper, R. J. (2014). The flipped classroom: A course redesign to foster learning and engagement in a health professions school. *Academic Medicine*, *89*(2), 236-243. doi:10.1097/ACM.0000000000000086
45. O'Flaherty, J., & Phillips, C. (2015). The use of flipped classrooms in higher education: A scoping review. *The internet and higher education*, *25*, 85-95.

46. Cook, D. A., & Ellaway, R. H. (2015). Evaluating technology-enhanced learning: a comprehensive framework. *Medical teacher*, *37*(10), 961-970.

47. Chmiel, A. S., Shaha, M., & Schneider, D. K. (2017). Introduction of blended learning in a master program: developing an integrative mixed method evaluation framework. *Nurse education today*, *48*, 172-179.

48. Menon, A., Korner-Bitensky, N., Chignell, M., & Straus, S. (2012). Usability testing of two e-learning resources: methods to maximize potential for clinician use. *Journal of rehabilitation medicine*, *44*(4), 338-345.

49. Nielsen, J. (2003). Usability 101: Introduction to usability.

50. Dubey, S. K., & Rana, A. (2010). Analytical roadmap to usability definitions and decompositions. *International Journal of Engineering Science and Technology*, *2*(9), 4723-4729.

51. Jeng, J. (2005). What is usability in the context of the digital library and how can it be measured?. *Information technology and libraries*, *24*(2), 3.

52. Phillips, R., Kennedy, G., & McNaught, C. (2012). The role of theory in learning technology evaluation research. *Australasian Journal of Educational Technology*, *28*(7), 1103-1118.

53. Taveira-Gomes, T., Saffarzadeh, A., Severo, M., Guimarães, M. J., & Ferreira, M. A. (2014). A novel collaborative e-learning platform for medical students-ALERT Student. *BMC medical education*, 14(1), 143.

54. Coughlan, J., & Morar, S. S. (2008). Development of a tool for evaluating multimedia for surgical education. *Journal of Surgical Research*, 149(1), 94-100.

55. Davids, M. R., Chikte, U. M., & Halperin, M. L. (2014). Effect of improving the usability of an e-learning resource: a randomized trial. *Advances in physiology education*, 38(2), 155-160.

56. Dow, A. W., DiazGranados, D., Mazmanian, P. E., & Retchin, S. M. (2014). An exploratory study of an assessment tool derived from the competencies of the interprofessional education collaborative. *Journal of interprofessional care*, *28*(4), 299-304.

57. te Pas, E., Wieringa-de Waard, M., Blok, B. S., Pouw, H., & van Dijk, N. (2016). Didactic and technical considerations when developing e-learning and CME. *Education and Information Technologies*, 21(5), 991-1005.

58. Walji, M. F., Kalenderian, E., Piotrowski, M., Tran, D., Kookal, K. K., Tokede, O., . . . Stark, P. C. (2014). Are three methods better than one? A comparative assessment of usability evaluation methods in an EHR. *International journal of medical informatics*, 83(5), 361-367.

59. Green, J., & Thorogood, N. (2013). *Qualitative methods for health research* (Third Edition): Sage.

60. DiCicco-Bloom, B., & Crabtree, B. F. (2006). The qualitative research interview. *Medical education*, *40*(4), 314-321.

61. Morgan, D. L. (1996). *Focus groups as qualitative research* (Vol. 16). Sage publications.

62. Nielsen, J. (1994, April). Usability inspection methods. In *Conference companion on Human factors in computing systems* (pp. 413-414). ACM.

63. Bevan, N. (2009). International standards for usability should be more widely used. *Journal of Usability Studies*, *4*(3), 106-113.

64. Arksey, H., & O'Malley, L. (2005). Scoping studies: towards a methodological framework. *International journal of social research methodology*, *8*(1), 19-32.

65. Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., ... & Hempel, S. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Annals of internal medicine*, *169*(7), 467-473.

66. Colquhoun, H. L., Levac, D., O'Brien, K. K., Straus, S., Tricco, A. C., Perrier, L., . . . Moher, D. (2014). Scoping reviews: time for clarity in definition, methods, and reporting. *Journal of clinical epidemiology*, *67*(12), 1291-1294.

67. Morris, M., Boruff, J. T., & Gore, G. C. (2016). Scoping reviews: establishing the role of the librarian. *Journal of the Medical Library Association: JMLA*, *104*(4), 346.

68. Peters, M. D., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., & Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *International journal of evidence-based healthcare*, *13*(3), 141-146.

69. Scopus: Access and use Support Centre. (2018). Elsevier. Retrieved from: https://service.elsevier.com/app/answers/detail/a_id/15534/supporthub/scopus/#tips

70. Eric. (2018). EBSCO. Retrieved from: https://www.ebsco.com/products/research-databases/eric

71. Mann, K., Gordon, J., & MacLeod, A. (2009). Reflection and reflective practice in health professions education: a systematic review. *Advances in health sciences education*, *14*(4), 595.

72. Definition and list of health professionals. (n.d.) Transformative Education for Health Professionals. Retrieved from: https://whoeducationguidelines.org/content/1-definition-and-list-health-professionals

73. Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative health research*, *15*(9), 1277-1288.

74. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77-101.

75. Kirkpatrick, D. L. (1979). Techniques for evaluating training programs. *Training and development journal.*

76. Praslova, L. (2010). Adaptation of Kirkpatrick's four level model of training criteria to assessment of learning outcomes and program evaluation in higher education. *Educational Assessment, Evaluation and Accountability*, 22(3), 215-225.

77. Ardito, C., Costabile, M. F., De Marsico, M., Lanzilotti, R., Levialdi, S., Roselli, T., & Rossano, V. (2006). An approach to usability evaluation of e-learning applications. *Universal access in the information society*, *4*(3), 270-283.

78. Benson, L., Elliott, D., Grant, M., Holschuh, D., Kim, B., Kim, H., ... & Reeves, T. C. (2002). Usability and instructional design heuristics for e-learning evaluation. In *EdMedia: World Conference on Educational Media and Technology* (pp. 1615-1621). Association for the Advancement of Computing in Education (AACE).

79. Costabile, M. F., De Marsico, M., Lanzilotti, R., Plantamura, V. L., & Roselli, T. (2005, January). On the usability evaluation of e-learning applications. In *null* (p. 6b). IEEE.

80. Rodríguez, C., Bartlett-Esquilant, G., Boillat, M., Dove, M., Grad, R., Lalla, L., ... & Bergman, H. (2015). Manifesto for family medicine educational research. *Canadian Family Physician*, *61*(9), 745-747.

81. Webster, F., Krueger, P., MacDonald, H., Archibald, D., Telner, D., Bytautas, J., & Whitehead, C. (2015). A scoping review of medical education research in family medicine. *BMC medical education*, *15*(1), 79.

82. Alonso-Sardón, M., Iglesias-de-Sena, H., Sáez-Lorenzo, M., Fernández, A. J. C., Salvat-Puig, J., & Mirón-Canelo, J. A. (2015). B-learning training in the certification of causes of death. *Journal of Forensic and Legal Medicine, 29*, 1-5. doi:10.1016/j.jflm.2014.10.003

83. Bohaty, B. S., Redford, G. J., & Gadbury-Amyot, C. C. (2016). Flipping the classroom: Assessment of strategies to promote student-centered, self-directed learning in a dental school course in pediatric dentistry. *Journal of Dental Education, 80*(11), 1319-1327.

84. Bonnes, S. L., Ratelle, J. T., Halvorsen, A. J., Carter, K. J., Hafdahl, L. T., Wang, A. T., . . . Wittich, C. M. (2017). Flipping the Quality Improvement Classroom in Residency Education. *Academic Medicine, 92*(1), 101-107. doi:10.1097/ACM.0000000000001412

85. Bösner, S., Pickert, J., & Stibane, T. (2015). Teaching differential diagnosis in primary care using an inverted classroom approach: Student satisfaction and gain in skills and knowledge. *BMC Medical Education, 15*(1). doi:10.1186/s12909-015-0346-x

86. Bossaer, J. B., Panus, P., Stewart, D. W., Hagemeier, N. E., & George, J. (2016). Student performance in a pharmacotherapy oncology module before and after flipping the classroom. *American Journal of Pharmaceutical Education, 80*(2). doi:10.5688/ajpe80231

87. Boysen-Osborn, M., Anderson, C. L., Navarro, R., Yanuck, J., Strom, S., McCoy, C. E., . . . Langdorf, M. I. (2016). Flipping the Advanced Cardiac Life Support Classroom with Team-based Learning: Comparison of Cognitive Testing Performance for Medical Students at the University of California, Irvine, United States. *Journal of educational evaluation for health professions, 13*, 11. doi:10.3352/jeehp.2016.13.11

88. Choi, H., Kim, J., Bang, K. S., Park, Y. H., Lee, N. J., & Kim, C. (2015). Applying the flipped learning model to an English-medium nursing course. *Journal of Korean Academy of Nursing, 45*(6), 939-948. doi:10.4040/jkan.2015.45.6.939

89. Crawford, J., Brudnoy, L., Soong, T., & Graham, T. (2013). Patient navigation in oncology nursing: An innovative blended learning model. *Journal of Continuing Education in Nursing, 44*(10), 461-469. doi:10.3928/00220124-20130903-88

90. Domínguez, L. C., Sierra, D., Pepín, J. J., Moros, G., & Villarraga, A. (2017). Effect of the Extended Inverted Classroom on clinical simulation for the resuscitation of trauma patients: Pilot study of student perceptions of learning. *Revista Colombiana de Anestesiologia, 45*, 4-11. doi:10.1016/j.rca.2017.07.011

91. Duque, G., Demontiero, O., Whereat, S., Gunawardene, P., Leung, O., Webster, P., . . . Sharma, A. (2013). Evaluation of a blended learning model in geriatric medicine: A successful learning experience for medical students. *Australasian Journal on Ageing, 32*(2), 103-109. doi:10.1111/j.1741-6612.2012.00620.x

92. Eachempati, P., Kiran Kumar, K., & Sumanth, K. (2016). Blended learning for reinforcing dental pharmacology in the clinical years: A qualitative analysis. *Indian Journal of Pharmacology, 48*(7), S25-S28. doi:10.4103/0253-7613.193315

93. Edwards, G., Kitzmiller, R. R., & Breckenridge-Sproat, S. (2012). Innovative health information technology training: Exploring blended learning. *CIN - Computers Informatics Nursing, 30*(2), 104-109. doi:10.1097/NCN.0b013e31822f7f7a

94. Elebiary, H. A., & Al Mahmoud, S. (2013). Enhancing blended courses to facilitate student achievement of learning outcomes. *Life Science Journal, 10*(2), 401-407.

95. Forrest, E. (2010). Perinatal mental health education for midwives in Scotland. *British Journal of Midwifery, 18*(5), 280-284. doi:10.12968/bjom.2010.18.5.47853

96. Furnes, M., Kvaal, K. S., & Høye, S. (2018). Communication in mental health nursing - Bachelor Students' appraisal of a blended learning training programme - An exploratory study. *BMC Nursing, 17*(1). doi:10.1186/s12912-018-0288-9

97. Gilboy, M. B., Heinerichs, S., & Pazzaglia, G. (2015). Enhancing student engagement using the flipped classroom. *Journal of Nutrition Education and Behavior, 47*(1), 109-114. doi:10.1016/j.jneb.2014.08.008

98. Goodie, J. L., Williams, P. M., Kurzweil, D., & Marcellas, K. B. (2011). Can blended classroom and distributed learning approaches be used to teach medical students how to initiate behavior change counseling during a clinical clerkship? *Journal of Clinical Psychology in Medical Settings, 18*(4), 353-360. doi:10.1007/s10880-011-9261-4

99. Gopalan, C., & Klann, M. C. (2017). The effect of flipped teaching combined with modified team-based learning on learner performance in physiology. *Advances in Physiology Education, 41*(3), 363-367. doi:10.1152/advan.00179.2016

100. Gostelow, N., Barber, J., Gishen, F., & Berlin, A. (2018). Flipping social determinants on its head: Medical student perspectives on the flipped classroom and simulated patients to teach social determinants of health. *Medical Teacher, 40*(7), 728-735. doi:10.1080/0142159X.2018.1436757

101. Green, R. A., & Whitburn, L. Y. (2016). Impact of introduction of blended learning in gross anatomy on student outcomes. *Anatomical Sciences Education, 9*(5), 422-430. doi:10.1002/ase.1602

102. Howlett, D., Vincent, T., Watson, G., Owens, E., Webb, R., Gainsborough, N., . . . Vincent, R. (2011). Blending online techniques with traditional face to face teaching methods to deliver

final year undergraduate radiology learning content. *European Journal of Radiology, 78*(3), 334-341. doi:10.1016/j.ejrad.2009.07.028

103. Kangwantas, K., Pongwecharak, J., Rungsardthong, K., Jantarathaneewat, K., Sappruetthikun, P., & Maluangnon, K. (2017). Implementing a flipped classroom approach to a course module in fundamental nutrition for pharmacy students. *Pharmacy Education, 17*(1), 329-334.

104. Koo, C. L., Demps, E. L., Farris, C., Bowman, J. D., Panahi, L., & Boyle, P. (2016). Impact of flipped classroom design on student performance and perceptions in a pharmacotherapy course. *American Journal of Pharmaceutical Education, 80*(2). doi:10.5688/ajpe80233

105. Kühl, S. J., Toberer, M., Keis, O., Tolks, D., Fischer, M. R., & Kühl, M. (2017). Concept and benefits of the inverted classroom method for a competency-based biochemistry course in the pre-clinical stage of a human medicine course of studies. *GMS Journal for Medical Education, 34*(3). doi:10.3205/zma001108

106. Langenau, E. E., Lee, R., & Fults, M., III. (2017). Blended learning educational format for third-year pediatrics clinical rotation. *Journal of the American Osteopathic Association, 117*(4), 234-243. doi:10.7556/jaoa.2017.041

107. Leikola, S. N. S., Tuomainen, L., Ovaskainen, H., Peura, S., Sevón-Vilkman, N., Tanskanen, P., & Airaksinen, M. S. A. (2009). Continuing education course to attain collaborative comprehensive medication review competencies. *American Journal of Pharmaceutical Education, 73*(6). doi:10.5688/aj7306108

108. Lorimer, J., & Hilliard, A. (2009). Incorporating learning technologies into undergraduate radiography education. *Radiography, 15*(3), 214-219. doi:10.1016/j.radi.2009.02.003

109. Makhdoom, N., Khoshhal, K. I., Algaidi, S., Heissam, K., & Zolaly, M. A. (2013). 'Blended learning' as an effective teaching and learning strategy in clinical medicine: A comparative cross-sectional university-based study. *Journal of Taibah University Medical Sciences, 8*(1), 12-17. doi:10.1016/j.jtumed.2013.01.002

110. Mary, S., Julie, J., & Jennifer, G. (2014). Teaching evidence based practice and research through blended learning to undergraduate midwifery students from a practice based perspective. *Nurse Education in Practice, 14*(2), 220-224. doi:10.1016/j.nepr.2013.10.001

111. Matsuda, Y., Azaiza, K., & Salani, D. (2017). Flipping the Classroom without Flipping Out the Students: Working with an Instructional Designer in an Undergraduate Evidence-Based Nursing Practice Course. *Quarterly Review of Distance Education, 18*(1), 17-27.

112. McLaughlin, J. E., Gharkholonarehe, N., Khanova, J., Deyo, Z. M., & Rodgers, J. E. (2015). Instructional design and assessment the impact of blended learning on student performance in a cardiovascular pharmacotherapy course. *American Journal of Pharmaceutical Education, 79*(2). doi:10.5688/ajpe79224

113. McLaughlin, J. E., Roth, M. T., Glatt, D. M., Gharkholonarehe, N., Davidson, C. A., Griffin, L. M., . . . Mumper, R. J. (2014). The flipped classroom: A course redesign to foster learning and engagement in a health professions school. *Academic Medicine, 89*(2), 236-243. doi:10.1097/ACM.0000000000000086

114. Milic, N. M., Trajkovic, G. Z., Bukumiric, Z. M., Cirkovic, A., Nikolic, I. M., Milin, J. S., . . . Stanisavljevic, D. M. (2016). Improving education in medical statistics: Implementing a blended learning model in the existing curriculum. *PloS One, 11*(2). doi:10.1371/journal.pone.0148882

115. Mohanna, K., Waters, M., & Deighan, M. (2008). Designing effective blended learning environments for training trainers in primary care. *Education for Primary Care, 19*(6), 597-604. doi:10.1080/14739879.2008.11493738

116. Moraros, J., Islam, A., Yu, S., Banow, R., & Schindelka, B. (2015). Flipping for success: Evaluating the effectiveness of a novel teaching approach in a graduate level setting. *BMC Medical Education, 15*(1). doi:10.1186/s12909-015-0317-2

117. Morton, D. A., & Colbert-Getz, J. M. (2017). Measuring the impact of the flipped anatomy classroom: The importance of categorizing an assessment by Bloom's taxonomy. *Anatomical Sciences Education, 10*(2), 170-175. doi:10.1002/ase.1635

118. Ocak, M. A., & Topal, A. D. (2015). Blended learning in anatomy education: A study investigating medical students' perceptions. *Eurasia Journal of Mathematics, Science and Technology Education, 11*(3), 647-683. doi:10.12973/eurasia.2015.1326a

119. O'Connor, E. E., Fried, J., McNulty, N., Shah, P., Hogg, J. P., Lewis, P., . . . Reddy, S. (2016). Flipping Radiology Education Right Side Up. *Academic Radiology, 23*(7), 810-822. doi:10.1016/j.acra.2016.02.011

120. Oh, J., Kim, S. J., Kim, S., & Vasuki, R. (2017). Evaluation of the effects of flipped learning of a nursing informatics course. *Journal of Nursing Education, 56*(8), 477-483. doi:10.3928/01484834-20170712-06

121. Ohtake, P. J., Lyons, A., Glogowski, M., Stellrecht, E., Aronoff, N., Grabowski, J., & Zafron, M. L. (2018). Using an interprofessional flipped classroom educational strategy for developing evidence-based practice knowledge and skills. *Journal of Interprofessional Education and Practice, 11*, 7-11. doi:10.1016/j.xjep.2017.12.010

122. Page, J., Meehan-Andrews, T., Weerakkody, N., Hughes, D. L., & Rathner, J. A. (2017). Student Perceptions and Learning Outcomes of Blended Learning in a Massive First-Year Core Physiology for Allied Health Subjects. *Advances in Physiology Education, 41*(1), 44-55.

123. Peisachovich, E. H., Murtha, S., Phillips, A., & Messinger, G. (2016). Flipping the Classroom: A Pedagogical Approach to Applying Clinical Judgment by Engaging, Interacting, and Collaborating with Nursing Students. *International Journal of Higher Education, 5*(4), 114-121.

124. Phillips, J. A., Schumacher, C., & Arif, S. (2016). Time spent, workload, and student and faculty perceptions in a blended learning environment. *American Journal of Pharmaceutical Education, 80*(6).

125. Popovic, N., Popovic, T., Dragovic, I. R., & Cmiljanic, O. (2018). A Moodle-based blended learning solution for physiology education in Montenegro: A case study. *Advances in Physiology Education, 42*(1), 111-117. doi:10.1152/ADVAN.00155.2017

126. Rose, E., Claudius, I., Tabatabai, R., Kearl, L., Behar, S., & Jhun, P. (2016). The Flipped Classroom in Emergency Medicine Using Online Videos with Interpolated Questions. *Journal of Emergency Medicine, 51*(3), 284-291.e281. doi:10.1016/j.jemermed.2016.05.033

127. Rose, E., Jhun, P., Baluzy, M., Hauck, A., Huang, J., Wagner, J., . . . Claudius, I. (2018). Flipping the classroom in medical student education: Does priming work? *Western Journal of Emergency Medicine, 19*(1), 93-100. doi:10.5811/westjem.2017.8.35162

128. Salajegheh, A., Jahangiri, A., Dolan-Evans, E., & Pakneshan, S. (2016). A combination of traditional learning and e-learning can be more effective on radiological interpretation skills in medical students: A pre- and post-intervention study Approaches to teaching and learning. *BMC Medical Education, 16*(1). doi:10.1186/s12909-016-0569-5

129. Stewart, A., Inglis, G., Jardine, L., Koorts, P., & Davies, M. W. (2013). A randomised controlled trial of blended learning to improve the newborn examination skills of medical students. *Archives of disease in childhood. Fetal and neonatal edition, 98*(2), F141-144. doi:10.1136/archdischild-2011-301252

130. Tang, F., Chen, C., Zhu, Y., Zuo, C., Zhong, Y., Wang, N., . . . Liang, D. (2017). Comparison between flipped classroom and lecture-based classroom in ophthalmology clerkship. *Medical Education Online, 22*(1), 1395679. doi:10.1080/10872981.2017.1395679

131. Ward, J. A., Beaton, R. D., Bruck, A. M., & De Castro, A. B. (2011). Promoting occupational health nursing training: An educational outreach with a blended model of distance and traditional learning approaches. *AAOHN Journal, 59*(9), 401-406. doi:10.3928/08910162-20110825-02

132. Wong, T. H., Ip, E. J., Lopes, I., & Rajagopalan, V. (2014). Pharmacy students' performance and perceptions in a flipped teaching pilot on cardiac arrhythmias. *American Journal of Pharmaceutical Education, 78*(10), 1-5.

133. Ye, L., & Smith, A. (2015). Developing and testing a sleep education program for college nursing students. *Journal of Nursing Education, 54*(9), 532-535. doi:10.3928/01484834-20150814-09

134. Zomorodi, M., de Saxe Zerden, L., Nance-Floyd, B., Alexander, L., Wilfert, R., & Byerley, J. (2017). Impact of an interprofessional population health course and clinical immersion experience: Students and practice outcomes. *Journal of Interprofessional Education and Practice, 9*, 91-94. doi:10.1016/j.xjep.2017.08.008

135. Ellner, A. L., & Phillips, R. S. (2017). The coming primary care revolution. *Journal of general internal medicine, 32*(4), 380-386.

136. Starfield, B., Shi, L., & Macinko, J. (2005). Contribution of primary care to health systems and health. *The milbank quarterly*, *83*(3), 457-502.

137. Carver T, Rodriguez C, Bergman H, Simmons R, Wang W. (2016) Assessing an International Medical Education Innovation: Blended Education Program for Teacher Leaders in General Practice in Shenzhen.

138. Carver T, Rodriguez C, Bergman H, Simmons R. (2017) Assessing an International Medical Education Innovation: Blended Education Program for Teacher Leaders at Hospital Santa Marcelina in Sao Paulo, Brazil

139. Carver T, Rodriguez C, Bergman H, Simmons R. (2018) Assessing an Innovative Blended Education Program: International Medical Education for General Practitioners

140. Carver T, Rodriguez C, Juras A, Bergman H, Simmons R. (2018) Assessing a Blended Education Program: Faculty Development Program on Fundamental Topics for Clinical Teachers at the Department of Family Medicine at McGill University

# Appendix

## Appendix 1: Search Strategy for Scopus

((TITLE-ABS-KEY(usability)) OR (TITLE-ABS-KEY("user experience")) OR (TITLE-ABS-KEY("human computer interaction")) OR (TITLE-ABS-KEY("user computer interface")) OR (TITLE-ABS-KEY(assess*)) OR (TITLE-ABS-KEY(test*)) OR (TITLE-ABS-KEY(satisfaction)) OR (TITLE-ABS-KEY(effective*)) OR (TITLE-ABS-KEY(efficiency)) OR (TITLE-ABS-KEY(efficacy)) OR (TITLE-ABS-KEY(learnability)) OR (TITLE-ABS-KEY(memorability)) OR (TITLE-ABS-KEY(evaluat*)) OR (TITLE-ABS-KEY(success W/4 program)))

AND ((TITLE-ABS-KEY("blended learning")) OR (TITLE-ABS-KEY(computer PRE/1 (learning OR instruction))) OR (TITLE-ABS-KEY("flipped learning")) OR (TITLE-ABS-KEY("flipped classroom")) OR (TITLE-ABS-KEY("inverted learning")) OR (TITLE-ABS-KEY("inverted classroom")) OR (TITLE-ABS-KEY("hybrid learning")) OR (TITLE-ABS-KEY("hybrid instruction")) OR (((TITLE-ABS-KEY("classroom-based learning")) AND TITLE-ABS-KEY("e-learning"))))

AND ((TITLE-ABS-KEY("medical education")) OR (TITLE-ABS-KEY("medical school")) OR (TITLE-ABS-KEY("medical curriculum")) OR (TITLE-ABS-KEY(medic*)) OR (TITLE-ABS-KEY("nurs*")) OR (TITLE-ABS-KEY("health profession*")) OR (TITLE-ABS-KEY("problem-based learning")) OR (TITLE-ABS-KEY("competency-based education")) OR (TITLE-ABS-KEY("competency-based learning")) OR (TITLE-ABS-KEY("clinical competence")) OR (TITLE-ABS-KEY(midwife*)) OR (TITLE-ABS-KEY(psych*)) OR (TITLE-ABS-KEY("mental health")) OR (TITLE-ABS-KEY(osteopath*)) OR (TITLE-ABS-KEY(naturopath*)) OR (TITLE-

ABS-KEY(homeopath*)) OR (TITLE-ABS-KEY(physio*)) OR (TITLE-ABS-KEY(pharmac*)) OR (TITLE-ABS-KEY(dent*)) OR (TITLE-ABS-KEY(chiropract*)) OR (TITLE-ABS-KEY(therap*)) OR (TITLE-ABS-KEY(assistant)) OR (TITLE-ABS-KEY("primary care")))

## Appendix 2: Search Strategy for ERIC (EBSCO)

S1 = usability OR user experience OR human computer interaction OR user computer interface OR assess* OR test* OR satisfaction OR effective* OR efficacy OR evaluat* OR efficiency OR learnability OR memorability OR (success W4 program)

S2 = blended learning OR ( computer N1 (learning OR instruction) ) OR computer uses in education OR flipped classroom OR flipped learning OR inverted classroom OR inverted learning OR hybrid learning OR hybrid instruction

S3 = (e-learning or online learning or web-based learning ) AND ( traditional learning or face-to-face )

S4 = S2 OR S3

S5 = (medical education or medical school or medical learners or medical curriculum or medical learner education or clinical education ) OR ( medicine or medical or health or healthcare ) OR ( nurse education or nursing education ) OR ( nurse or nurses or nursing or nursing staff or health care professional ) OR professional education OR allied health professions OR allied health occupations education OR problem based learning OR competency-based education OR competency based learning OR clinical competence OR clinical experience OR clinical teaching

S6 = clinical education OR midwi* OR psych* OR mental health or mental illness or mental disorder or psychiatric illness OR osteopath* OR naturopath* OR homeopath* OR physio* OR pharmac* OR therap* OR dent* OR chiropract*

S7 = assistant OR primary care or primary health care or primary healthcare OR health occupations OR health personnel OR health sciences

S8 = S5 OR S6 OR S7

Final Search Strategy S9 = S1 AND S4 AND S8

## Appendix 3: Title, Abstract, & Full-Text Screening Questionnaire

1. Is the study about a program for health professions education?
    a. Yes → According to the inclusion/exclusion criteria, this would constitute a health professional (e.g. nurse, doctor, physiotherapist) or learner in a health professions program (e.g. nursing program, medical school, physiotherapy program) that functions to treat humans
    b. No → According to the inclusion/exclusion criteria, a study that targets its BLP towards veterinary learners for example will be excluded from the review
    c. Maybe → The text (title, abstract, or full-text) does not clearly indicate who the BLP is targeted to
2. Does the study assess a blended learning program (includes both in-person synchronous learning components and asynchronous online learning environment that can be accessed conveniently/at a learner's own time)?
    a. Yes
    b. No
    c. Maybe
3. Does the BLP utilize asynchronous online learning methods for at least 30 to 80% of the program?
    a. Yes
    b. No
    c. Maybe
4. Does the BLP make use of a learning management system or indicate that learner use of online learning platforms was tracked?

a. Yes

b. No

c. Maybe

5. Does the study evaluate for usability or one of its main components (effectiveness, satisfaction, or efficiency)?

    a. Yes

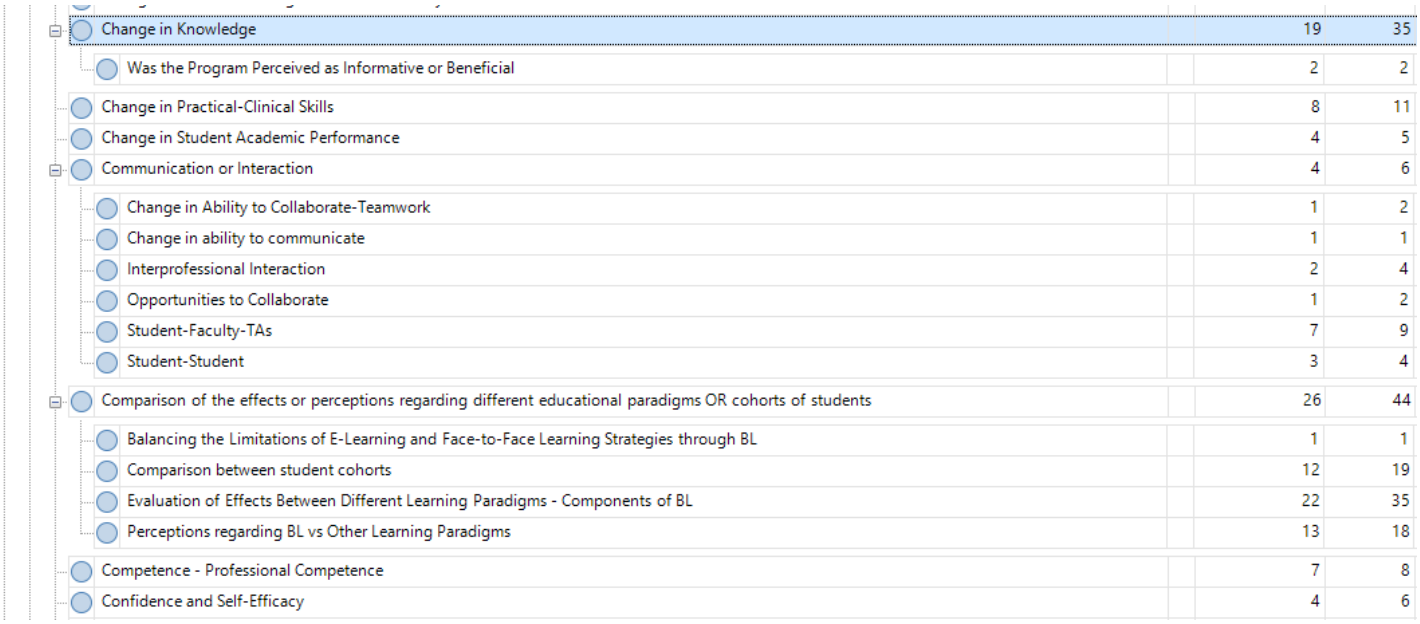
    b. No

    c. Maybe

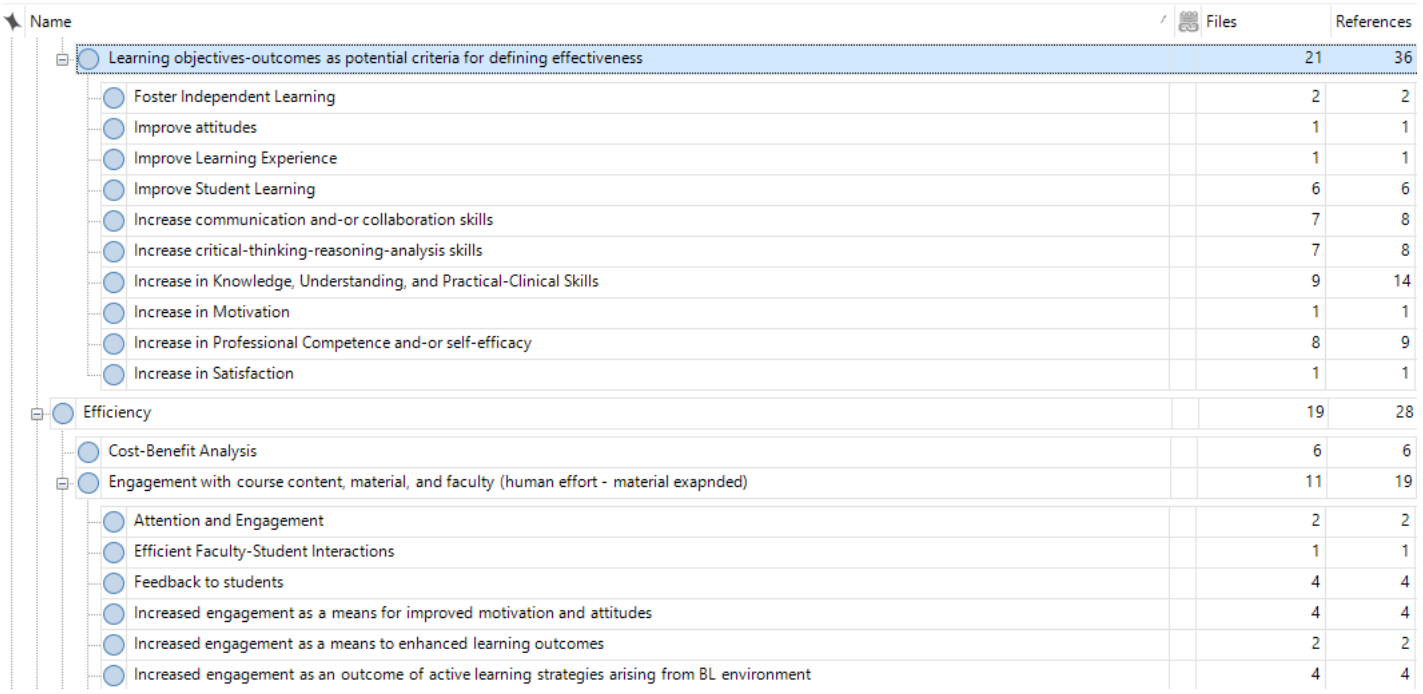
# Appendix 4: NVIVO Codes for Qualitative Analysis

| Name | Files | Reference |
|---|---|---|
| 1- Avoiding the 'Usability' Label and Using Undefined Related Terms | 0 | 0 |
| Frequently Applied Ambiguous Terms Used to Discuss Effects of BLPs Aside from Effectiveness, Efficiency, Satisfaction, Perceptions, Accessibility | 0 | 0 |
| Accessibility | 15 | 26 |
| Appropriateness | 16 | 19 |
| Clarity | 12 | 18 |
| Comprehensiveness | 1 | 1 |
| Convenience | 3 | 4 |
| Ease of Use | 9 | 13 |
| Efficacy | 4 | 5 |
| Favourable | 1 | 1 |
| Helpful | 5 | 6 |
| Informative | 1 | 1 |
| Perceived value | 1 | 1 |
| Relevance | 10 | 14 |
| Use-Usefulness | 25 | 34 |
| Utility | 1 | 1 |
| Worthwhile | 1 | 1 |

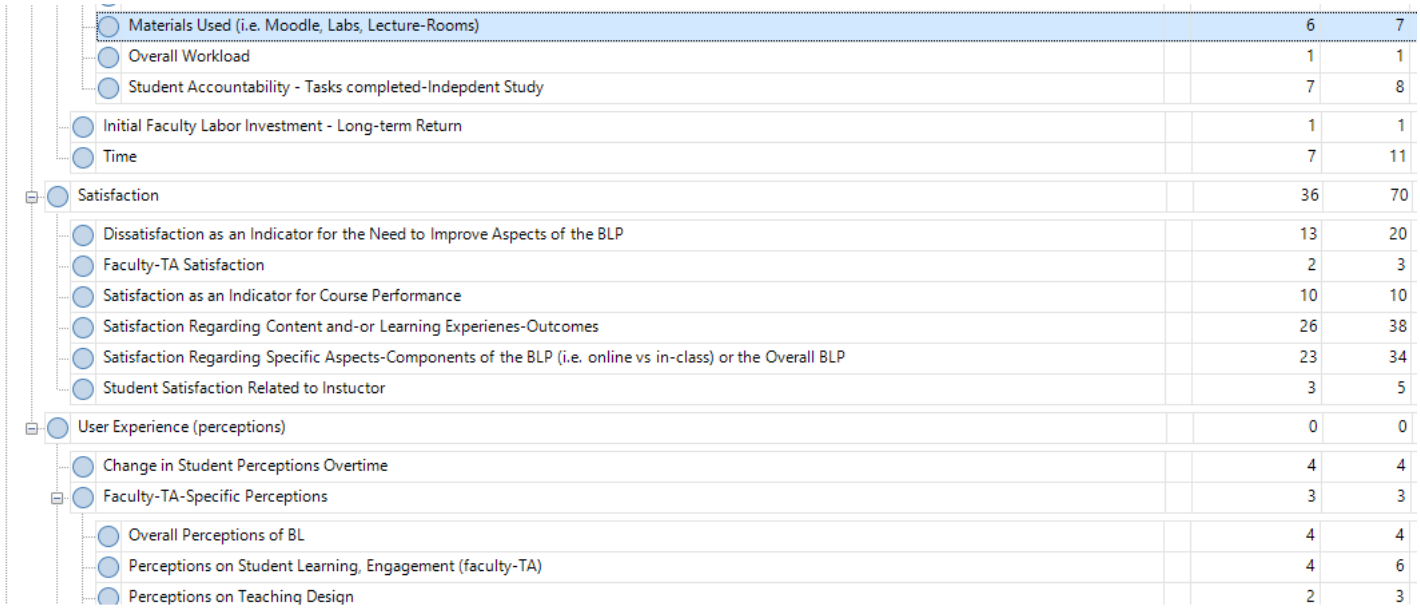
| Name | Files | References |
|---|---|---|
| 2- Confusing Conceptualization of the Components of Usability (i.e. Effectiveness, Efficiency, and Satisfaction) | 0 | 0 |
| Interrelation of Usability Concepts and Related Terminology | 0 | 0 |
| Blending of Effectiveness, Efficiency, Satisfaction, and Perceptions | 3 | 4 |
| Effciency and Satisfaction - Discussed As Related | 3 | 3 |
| Effectiveness or Efficiency - The Unclear Devision Between These Terms | 16 | 20 |
| Satisfaction and Effectiveness - Heavily Related, Sometimes Borderline Inseperable | 13 | 18 |
| The unclear devision between Satisfaction and Perceptions | 7 | 7 |
| Key Dimensions of Usability and Related Concepts that are Consistently Assessed in BLPs | 0 | 0 |
| Accessibility | 2 | 3 |
| Ability to access material regardless of time and place - flexibility | 8 | 15 |
| Access as a Means of Measuring Efficiency or Effectiveness | 2 | 3 |
| Easy access | 5 | 6 |
| Effectiveness | 29 | 41 |
| Appropriateness of Content and Course Materials | 1 | 1 |
| Attitudes, Behaviour, Values | 6 | 10 |
| Change in Motivation | 5 | 7 |
| Increased Control - Self-Regulated Learning - Autonomy | 5 | 6 |
| Benefits-Disadvantages of E-Learning or In-Person Aspects of BL (Avoidance of Harm from Use) | 14 | 16 |
| Change in Decision Making and-or Critical Analysis Skills | 7 | 10 |



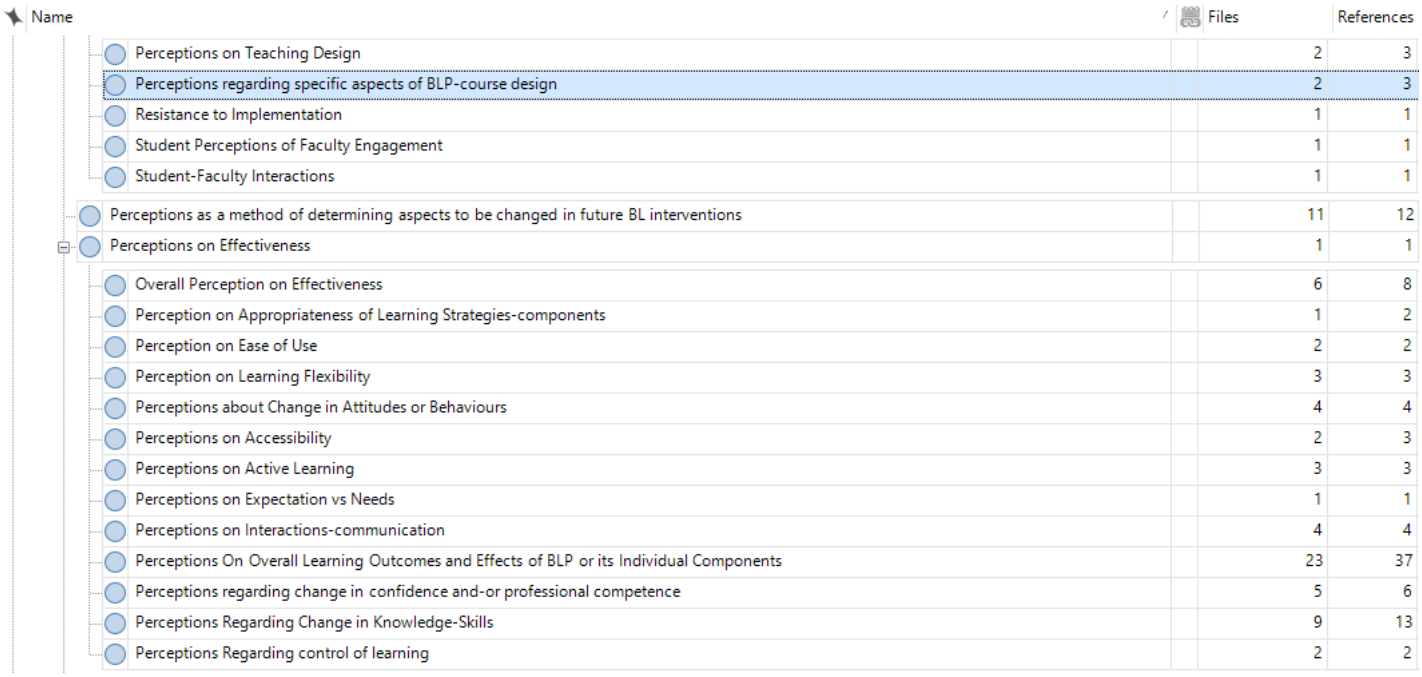
| | | |
|---|---|---|
| Change in Knowledge | 19 | 35 |
| Was the Program Perceived as Informative or Beneficial | 2 | 2 |
| Change in Practical-Clinical Skills | 8 | 11 |
| Change in Student Academic Performance | 4 | 5 |
| Communication or Interaction | 4 | 6 |
| Change in Ability to Collaborate-Teamwork | 1 | 2 |
| Change in ability to communicate | 1 | 1 |
| Interprofessional Interaction | 2 | 4 |
| Opportunities to Collaborate | 1 | 2 |
| Student-Faculty-TAs | 7 | 9 |
| Student-Student | 3 | 4 |
| Comparison of the effects or perceptions regarding different educational paradigms OR cohorts of students | 26 | 44 |
| Balancing the Limitations of E-Learning and Face-to-Face Learning Strategies through BL | 1 | 1 |
| Comparison between student cohorts | 12 | 19 |
| Evaluation of Effects Between Different Learning Paradigms - Components of BL | 22 | 35 |
| Perceptions regarding BL vs Other Learning Paradigms | 13 | 18 |
| Competence - Professional Competence | 7 | 8 |
| Confidence and Self-Efficacy | 4 | 6 |



| Name | Files | References |
|---|---|---|
| Learning objectives-outcomes as potential criteria for defining effectiveness | 21 | 36 |
| Foster Independent Learning | 2 | 2 |
| Improve attitudes | 1 | 1 |
| Improve Learning Experience | 1 | 1 |
| Improve Student Learning | 6 | 6 |
| Increase communication and-or collaboration skills | 7 | 8 |
| Increase critical-thinking-reasoning-analysis skills | 7 | 8 |
| Increase in Knowledge, Understanding, and Practical-Clinical Skills | 9 | 14 |
| Increase in Motivation | 1 | 1 |
| Increase in Professional Competence and-or self-efficacy | 8 | 9 |
| Increase in Satisfaction | 1 | 1 |
| Efficiency | 19 | 28 |
| Cost-Benefit Analysis | 6 | 6 |
| Engagement with course content, material, and faculty (human effort - material exapnded) | 11 | 19 |
| Attention and Engagement | 2 | 2 |
| Efficient Faculty-Student Interactions | 1 | 1 |
| Feedback to students | 4 | 4 |
| Increased engagement as a means for improved motivation and attitudes | 4 | 4 |
| Increased engagement as a means to enhanced learning outcomes | 2 | 2 |
| Increased engagement as an outcome of active learning strategies arising from BL environment | 4 | 4 |



| | | |
|---|---|---|
| Materials Used (i.e. Moodle, Labs, Lecture-Rooms) | 6 | 7 |
| Overall Workload | 1 | 1 |
| Student Accountability - Tasks completed-Indepdent Study | 7 | 8 |
| Initial Faculty Labor Investment - Long-term Return | 1 | 1 |
| Time | 7 | 11 |
| Satisfaction | 36 | 70 |
| Dissatisfaction as an Indicator for the Need to Improve Aspects of the BLP | 13 | 20 |
| Faculty-TA Satisfaction | 2 | 3 |
| Satisfaction as an Indicator for Course Performance | 10 | 10 |
| Satisfaction Regarding Content and-or Learning Experienes-Outcomes | 26 | 38 |
| Satisfaction Regarding Specific Aspects-Components of the BLP (i.e. online vs in-class) or the Overall BLP | 23 | 34 |
| Student Satisfaction Related to Instuctor | 3 | 5 |
| User Experience (perceptions) | 0 | 0 |
| Change in Student Perceptions Overtime | 4 | 4 |
| Faculty-TA-Specific Perceptions | 3 | 3 |
| Overall Perceptions of BL | 4 | 4 |
| Perceptions on Student Learning, Engagement (faculty-TA) | 4 | 6 |
| Perceptions on Teaching Design | 2 | 3 |



| Name | Files | References |
|---|---|---|
| Perceptions on Teaching Design | 2 | 3 |
| Perceptions regarding specific aspects of BLP-course design | 2 | 3 |
| Resistance to Implementation | 1 | 1 |
| Student Perceptions of Faculty Engagement | 1 | 1 |
| Student-Faculty Interactions | 1 | 1 |
| Perceptions as a method of determining aspects to be changed in future BL interventions | 11 | 12 |
| Perceptions on Effectiveness | 1 | 1 |
| Overall Perception on Effectiveness | 6 | 8 |
| Perception on Appropriateness of Learning Strategies-components | 1 | 2 |
| Perception on Ease of Use | 2 | 2 |
| Perception on Learning Flexibility | 3 | 3 |
| Perceptions about Change in Attitudes or Behaviours | 4 | 4 |
| Perceptions on Accessibility | 2 | 3 |
| Perceptions on Active Learning | 3 | 3 |
| Perceptions on Expectation vs Needs | 1 | 1 |
| Perceptions on Interactions-communication | 4 | 4 |
| Perceptions On Overall Learning Outcomes and Effects of BLP or its Individual Components | 23 | 37 |
| Perceptions regarding change in confidence and-or professional competence | 5 | 6 |
| Perceptions Regarding Change in Knowledge-Skills | 9 | 13 |
| Perceptions Regarding control of learning | 2 | 2 |

| | | |
|---|---|---|
| Perceptions on Efficiency | 2 | 2 |
| In-class vs Online Activites | 10 | 12 |
| Perceptions Regarding Certain Materials Provided | 9 | 15 |
| Perceptions regarding student-faculty interactions | 3 | 3 |
| Perceptions related to contact-time | 1 | 1 |
| Perceptions Related to Effort Expended By Students or Faculty - Workload - By students and-or faculty | 13 | 16 |
| Perceptions related to the benefits or disadvantages of certain tasks | 5 | 5 |
| Perceptions related to the necessity of certain tasks-components of BLPs | 1 | 1 |
| Perceptions Related to Time | 9 | 11 |
| Perceptions on Overall Learning Experiences-Interatction with BLP - satisafction | 23 | 31 |
| The Discussed Benefits of Blended Learning - Implicit References to Usability Evaluations | 11 | 15 |
| BL as a means of applying active learning strategies | 8 | 11 |
| Improved teaching efficiency | 4 | 4 |
| Increased Engagement | 4 | 4 |
| Increased Feedback to Students and Instructors | 2 | 2 |
| Increased Higher-Order Thinking and Problem Solving Skills | 3 | 4 |
| Increased Knowledge Retention | 1 | 1 |
| Increased Learning Outcomes - Student Performance | 4 | 4 |

| | | |
|---|---|---|
| Increased Motivation and Attitudes | 2 | 2 |
| BL environment's ability to capture attention and-or increase student engagement | 4 | 6 |
| Catering to the audience - Younger Generations have Grown Up with Technology | 7 | 9 |
| Decreased overall cost to implement program | 4 | 5 |
| Enables Competency Acquision | 1 | 1 |
| Faciliation of Deeper Learning | 1 | 1 |
| Flexibility in Learning - Perceived Convenience | 4 | 8 |
| Greater Control-Ownership of Learning for Students | 6 | 7 |
| Improved teaching consistency across different cores-sites (i.e. Standardized content) | 1 | 2 |
| Increased opportunity to implement a wider range of learning styles through blending multiple pedagogies | 12 | 20 |
| Reduced burden on faculty in the long term | 1 | 2 |
| Usability important for in-depth learning and retention | 19 | 24 |

| | | |
|---|---|---|
| 3- Lack of Consensual Approach to Assessment | 0 | 0 |
| Variability in the Adoptions of Research Approaches to Evaluation | 27 | 46 |
| Differences in Researcher Values | 0 | 0 |
| Ethical concerns about evaluation or program pilot | 1 | 1 |
| Importance of applying specific evaluation model to improve quality of evaluation | 1 | 1 |
| Importance of assessing both pedagogy and technology | 2 | 2 |
| Importance of balancing in class and online learning methods | 1 | 1 |
| Importance of Control Groups | 1 | 1 |
| Importance of measuring satisfaction | 5 | 6 |
| Need for evidence-based approach | 1 | 1 |
| The need to Measure long-term changes | 7 | 9 |
| The need to utilize multiple strategies to evaluate effects of BLPs | 4 | 5 |
| Various Research Designs Used to Evaluate BLPs | 31 | 51 |
| Techniques to ensure validity | 10 | 12 |
| The Application of Both Objective and Subjective Measures | 32 | 52 |
| Different BLP Evaluations and Study Designs | 49 | 104 |

# Appendix 5: PRISMA-ScR Checklist

*Table.* PRISMA-ScR Checklist

| Section | Item | PRISMA-ScR Checklist Item |
|---|---|---|
| **Title** | 1 | Identify the report as a scoping review. |
| **Abstract** | | |
| Structured summary | 2 | Provide a structured summary that includes (as applicable) background, objectives, eligibility criteria, sources of evidence, charting methods, results, and conclusions that relate to the review questions and objectives. |
| **Introduction** | | |
| Rationale | 3 | Describe the rationale for the review in the context of what is already known. Explain why the review questions/objectives lend themselves to a scoping review approach. |
| Objectives | 4 | Provide an explicit statement of the questions and objectives being addressed with reference to their key elements (e.g., population or participants, concepts, and context) or other relevant key elements used to conceptualize the review questions and/or objectives. |
| **Methods** | | |
| Protocol and registration | 5 | Indicate whether a review protocol exists; state if and where it can be accessed (e.g., a Web address); and if available, provide registration information, including the registration number. |
| Eligibility criteria | 6 | Specify characteristics of the sources of evidence used as eligibility criteria (e.g., years considered, language, and publication status), and provide a rationale. |
| Information sources* | 7 | Describe all information sources in the search (e.g., databases with dates of coverage and contact with authors to identify additional sources), as well as the date the most recent search was executed. |
| Search | 8 | Present the full electronic search strategy for at least 1 database, including any limits used, such that it could be repeated. |
| Selection of sources of evidence† | 9 | State the process for selecting sources of evidence (i.e., screening and eligibility) included in the scoping review. |
| Data charting process‡ | 10 | Describe the methods of charting data from the included sources of evidence (e.g., calibrated forms or forms that have been tested by the team before their use, and whether data charting was done independently or in duplicate) and any processes for obtaining and confirming data from investigators. |
| Data items | 11 | List and define all variables for which data were sought and any assumptions and simplifications made. |
| Critical appraisal of individual sources of evidence§ | 12 | If done, provide a rationale for conducting a critical appraisal of included sources of evidence; describe the methods used and how this information was used in any data synthesis (if appropriate). |
| Summary measures | 13 | Not applicable for scoping reviews. |
| Synthesis of results | 14 | Describe the methods of handling and summarizing the data that were charted. |
| Risk of bias across studies | 15 | Not applicable for scoping reviews. |
| Additional analyses | 16 | Not applicable for scoping reviews. |
| **Results** | | |
| Selection of sources of evidence | 17 | Give numbers of sources of evidence screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally using a flow diagram. |
| Characteristics of sources of evidence | 18 | For each source of evidence, present characteristics for which data were charted and provide the citations. |
| Critical appraisal within sources of evidence | 19 | If done, present data on critical appraisal of included sources of evidence (see item 12). |
| Results of individual sources of evidence | 20 | For each included source of evidence, present the relevant data that were charted that relate to the review questions and objectives. |
| Synthesis of results | 21 | Summarize and/or present the charting results as they relate to the review questions and objectives. |
| Risk of bias across studies | 22 | Not applicable for scoping reviews. |
| Additional analyses | 23 | Not applicable for scoping reviews. |
| **Discussion** | | |
| Summary of evidence | 24 | Summarize the main results (including an overview of concepts, themes, and types of evidence available), link to the review questions and objectives, and consider the relevance to key groups. |
| Limitations | 25 | Discuss the limitations of the scoping review process. |
| Conclusions | 26 | Provide a general interpretation of the results with respect to the review questions and objectives, as well as potential implications and/or next steps. |
| **Funding** | 27 | Describe sources of funding for the included sources of evidence, as well as sources of funding for the scoping review. Describe the role of the funders of the scoping review. |

JBI = Joanna Briggs Institute; PRISMA-ScR = Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews.
* Where *sources of evidence* (see second footnote) are compiled from, such as bibliographic databases, social media platforms, and Web sites.
† A more inclusive/heterogeneous term used to account for the different types of evidence or data sources (e.g., quantitative and/or qualitative research, expert opinion, and policy documents) that may be eligible in a scoping review as opposed to only studies. This is not to be confused with *information sources* (see first footnote).
‡ The frameworks by Arksey and O'Malley (6) and Levac and colleagues (7) and the JBI guidance (4, 5) refer to the process of data extraction in a scoping review as data charting.
§ The process of systematically examining research evidence to assess its validity, results, and relevance before using it to inform a decision. This term is used for items 12 and 19 instead of "risk of bias" (which is more applicable to systematic reviews of interventions) to include and acknowledge the various sources of evidence that may be used in a scoping review (e.g., quantitative and/or qualitative research, expert opinion, and policy documents).

## Appendix 6: Institutional Research Ethics Board Approval

McGill

Faculty of Medicine
3655 Promenade Sir William Osler #633
Montreal, QC H3G 1Y6

Faculté de médecine
3655, Promenade Sir William Osler #633
Montréal, QC H3G 1Y6

Fax/Télécopieur: (514) 398-3870
Tél/Tel: (514) 398-3124

01 June 2018

Dr. Charo Rodriguez
Department of Family Medicine
5858 Ch Côte-des-Neiges, 3rd Floor, Room 328
Montreal QC H3S 1Z1

**RE: IRB Study Number A06-E42-18A**
*Developing and initially validating a new tool for measuring usability of an international blended learning initiative in Family Medicine Faculty Development: a mixed methods study*

Dear Dr. Rodriguez,

Thank you for submitting the above-referenced study for an ethics review. This study was reviewed on behalf of your Masters Student, Anish Arora.

As this study involves no more than minimal risk, and in accordance with Articles 2.9 and 6.12 of the 2014 Edition of the Canadian Tri-Council Policy Statement of Ethical Conduct for Research Involving Humans (TCPS2 2014) and U.S. Title 45 CFR 46, Section 110 (b), paragraph (1), we are pleased to inform you that an expedited approval for the above-referenced study (Study Protocol, and Consent Form, v. May 2018) was provided by the IRB Chair on 01 June 2018. The ethics certificate is valid until **31 May 2019**. The study proposal will be presented for corroborative approval at the next scheduled meeting of the Institutional Review Board, and a certification document will be issued to you at that time.

***Please submit a copy of the questionnaire to the McGill IRB once available.***

A review of all research involving human subjects is required on an annual basis in accord with the date of initial approval. The annual review should be submitted at least one month before **31 May 2019**. Please inform the IRB promptly of any modifications that may occur to the study over the next twelve months.

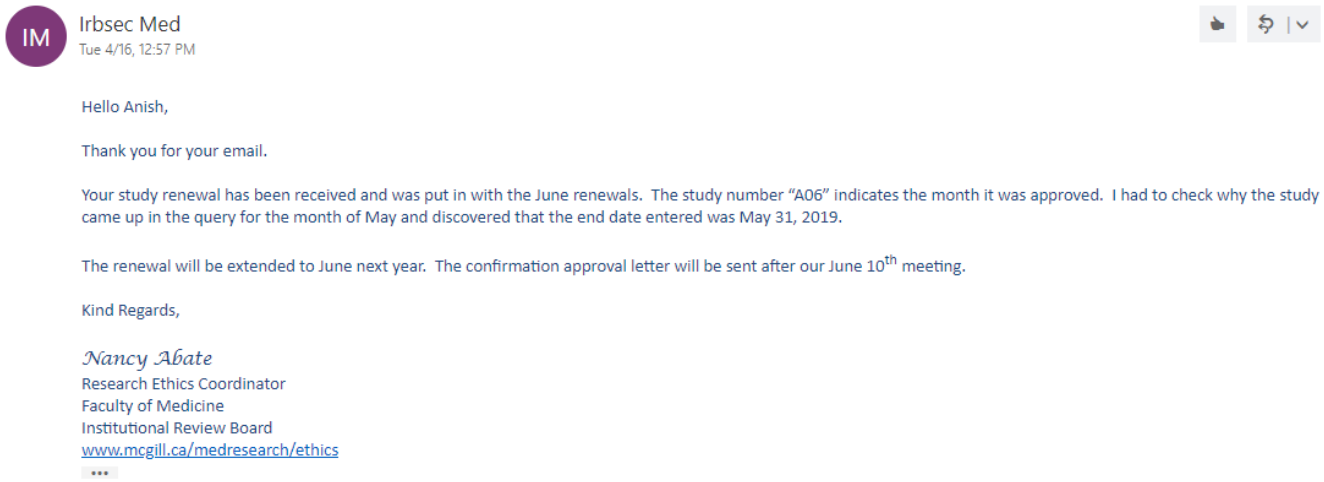
Sincerely,

Roberta Palmour

Roberta Palmour, PhD
Chair
Institutional Review Board

cc: Anish Arora
A06-E42-18A

## Appendix 7: Institutional Research Ethics Board – Continuing Review Acceptance



IM

Irbsec Med

Tue 4/16, 12:57 PM

Hello Anish,

Thank you for your email.

Your study renewal has been received and was put in with the June renewals. The study number "A06" indicates the month it was approved. I had to check why the study came up in the query for the month of May and discovered that the end date entered was May 31, 2019.

The renewal will be extended to June next year. The confirmation approval letter will be sent after our June 10th meeting.

Kind Regards,

*Nancy Abate*
Research Ethics Coordinator
Faculty of Medicine
Institutional Review Board
www.mcgill.ca/medresearch/ethics