A Protocol for a Systematic Review and Meta-Analysis on Social Presence

Dr. David Mykota

Abstract: The purpose of this research is to report the protocol for a systematic review and meta-analysis of the construct of social presence. Protocols for systematic reviews and meta-analysis are desirable as they offer a more knowledgeable and transparent conceptualization of a study so methods can be appraised and suggestions for modifications provided. The focused purpose outlined in the protocol is to systematically review and statistically summarize, thru meta-analysis (i.e., the study of studies), the research literature on social presence within higher education. The objectives of the systematic review are i) to determine the holistic effects that social presence has on student learning and student satisfaction outcomes across studies, and ii) what study characteristics (i.e., course design elements, instructor behaviors', and online learning environments) explain and moderate the variability in the results. The study protocol outlines the aims, objectives, and methods used to conduct the study and collect and analyze the data. The publication of the research protocol is an accountable process that mitigates research duplication while allowing for a more informed understanding of the prospective program of research. As such, with the completion of the systematic review and meta-analysis, the higher education online community will benefit from a more strategic and knowledgeable conceptualization of what constitutes social presence and how best to incorporate it into quality online learning environments.
Keywords: study protocol, evidence synthesis, higher education, online learning, systematic review, meta-analysis, course design, teaching, technology, social presence
Protocole de Revue Systématique et Méta-Analyse sur la Présence Sociale

Résumé: L'objectif de cette recherche est de présenter le protocole d’une revue systématique et d’une méta-analyse sur le concept de présence sociale. Les protocoles pour les revues systématiques et les méta-analyses sont intéressants en ce sens qu’ils offrent une conceptualisation plus éclairée et transparente d’une étude, ce qui permet d’en évaluer la méthode, voire de proposer des modifications. Le protocole décrit dans cet article a pour but d’examiner systématiquement et de résumer statistiquement, par le biais d’une méta-analyse (c’est-à-dire l’étude d’études), la littérature scientifique sur la présence sociale dans l’enseignement supérieur. Les objectifs de la revue systématique sont i) de déterminer les effets holistiques de la présence sociale sur les résultats de l’apprentissage et de la satisfaction des étudiants à travers les études, et ii) de définir les caractéristiques de l’étude (c.-à-d. les éléments de conception du cours, les comportements de l’instructeur et les environnements d’apprentissage en ligne) qui expliquent et modèrent la variabilité des résultats. Le protocole de recherche décrit ainsi les buts, les objectifs et les méthodes utilisées pour mener l’étude, recueillir et analyser les données. La publication du protocole de recherche est un processus responsable qui limite la duplication des recherches tout en permettant une compréhension plus éclairée du programme de recherche envisagé. Ainsi, avec cette publication de la revue systématique et de la méta-analyse, la communauté de l’enseignement supérieur en ligne bénéficiera d’une conceptualisation plus stratégique et mieux informée de ce qui constitue la présence sociale et de la meilleure façon de l’intégrer dans des environnements d’apprentissage en ligne de qualité.

Mots-clés: protocole d’étude, synthèse des données, enseignement supérieur, apprentissage en ligne, revue systématique, méta-analyse, conception de cours, enseignement, technologie, présence sociale.
Introduction

Systematic reviews are used to synthesize research literature in a rigorous and transparent manner. They are becoming more popular within the social sciences and in particular the field of education and eLearning (Caskurlu et al., 2020; Chapman, 2021; Petticrew and Roberts, 2006; Richardson et al., 2017; Zawacki-Richter et al., 2020). Systematic reviews are different from literature and narrative reviews as they include a series of distinct methodological stages that outline the research question, criteria for inclusion and exclusion of studies, the search strategy for selecting studies, method of data extraction, assessment of study quality, data analysis, and interpretation (Petticrew and Roberts, 2006; Uman, 2011; Zawacki-Richter et al., 2020). As such, systematic reviews are necessary and valuable because they allow for an objective understanding of the phenomenon being investigated; however, they are only one form of many different forms that reviews can be undertaken, each with their own advantages (Goagoses and Koglin, 2020; Pettigrew and Roberts, 2006).

Historically, systematic reviews were developed, in part, out of the need to inform evidenced based practices in the health sciences (Grant and Booth, 2009). These health sciences related reviews were often characterized by research involving randomized control trials and followed a PICO framework (i.e., population, intervention, comparator, and outcome) to determine the effectiveness of a treatment/intervention on a given population (Schardt et al., 2007). Although this approach proved useful to the understanding of treatment efficacy in the health sciences and medicine it was difficult to adapt to social sciences generally, and more specifically educational research. As research overviews in education evolved, they began to incorporate the methodology and reporting standards offered by systematic reviews (Polanin et al., 2017). Chapman (2021), in describing the characteristics of systematic reviews in the social sciences, found that there were some
methodological shortcomings in the way database searches were conducted, the number of reviewers involved in those searches, and the use of appropriate reporting standards that followed and adapted the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; https://prisma-statement.org/) checklist as developed by health sciences researchers (Moher et al., 2009). Nonetheless, according to Chapman (2021) in most of the cases reviewed the subject categories of education, psychology, and sociology tended to fair better in their methodological approaches and reporting standards. However, a clear and articulate published research protocol (Chapman, 2021) could easily have addressed some of the methodological concerns raised.

A published systematic review protocol is desirable because it is an open process by which methods can be appraised and duplication of research avoided. Published protocols can also aid in the detection of selective reporting while helping guide analysis and future research (Chapman, 2021; Shamseer et al., 2015; Tai et al., 2020). In the health sciences for example, publishing protocols for systematic reviews is encouraged to solicit feedback and reduce researcher bias, thereby improving the quality of the investigation being undertaken (Silagy et al.; 2002). Although there are a growing body of systematic reviews in the social sciences and education literature, few research protocols are being published or registered, especially in the field of eLearning and distance education. Subsequently, the International Database of Education Systematic Reviews (https://idesr.org/), with the assistance of Oxford University, was established to address this gap. The purpose of this brief review is to iterate the research protocol for a systematic review and meta-analysis on the construct social presence as applied to higher education online environs. In turn, a more open, transparent, and accountable process to improve the reporting quality of the proposed systematic review can potentially be obtained.
Overview

The social and interpersonal interactions that characterize the way individuals communicate, interrelate, and project themselves online is often described as social presence (Mykota, 2018). Since social interactions in education are an essential aspect of learning (Dewey, 1963; Hiltz, 1994; Hurst et al., 2013; Liaw and Huang, 2000) it is important to understand how social presence can be facilitated as it can impact the effectiveness of the online learning experience (Borup et al., 2012; Kim et al., 2011; Poth, 2018; Richardson and Swan, 2003). For example, social presence has been found to influence both student learning and academic performance (Hostteter and Bush, 2013; Joksimović et al., 2015; Mykota, 2018; Richardson et al.; 2017; Wise et al., 2004), along with student satisfaction of the online learning experience (Moallem, 2015; So and Brush, 2008)—while also improving student retention (Bowers and Kumar, 2015; Robb and Sutton, 2014).

Social presence has evolved over the years in concert with the development of the online learning medium. As online learning advanced, social presence theory’s conception of digital learning spaces progressed from strictly technologically computer mediated communication to socially and technologically co-determined environments (Gunawardena and Zittle, 1997; Tu and McIsaac, 2002). Initially conceived by social psychologists as comprising elements of both immediacy (Wiener and Mehrabian, 1968) and intimacy (Argyle and Dean, 1965), Short, et al., (1976) then applied these behaviours to the social psychology of telecommunications. Subsequently, Gunawardena and Zittle, (1997) along with Tu and McIsaac (2002), reconceptualized social presence as being co-determined by personal and social interactions within an educational context. With the advent of the Community of Inquiry (CoI) model, social presence theory came to recognize the importance of the group and community (Garrison, 2009), as well as the importance of teaching and cognitive presence. Gradually, social presence evolved from a two-
dimensional construct to one that is multidimensional involving a pattern of performable behaviours (e.g., Kehrwald, 2010; Kreijins et al., 2011; Kreijins et al., 2014; Sung and Mayer, 2012). Whiteside (2017), drawing on CoI research, integrated Vygotsky’s (1978) social development theory which views social presence as a critical literacy—thus offering a different perspective on social presence than initially proposed by Short et al. (1976). To help clarify our understanding of the construct, Krejins et al. (2021) identified social space and sociability as being linked to social presence, yet different in their conceptualization. Accordingly, we can understand and clarify some of the definitional and measurement issues that confound social presence research when we disentangle and differentiate these terms and discuss social presence as a unique psychological phenomenon (Kreijins et al., 2021).

To aid in the development of social presence, a scoping study identified that both course design and instructor behaviors can influence and augment the development of social presence and thereby effect student learning and satisfaction (Mykota, 2018). However, these findings were specific to independent individual studies, with the effect sizes varying across studies. The combined effects, statistically summarized, were not determined. The statistical aggregation of research findings (a meta-analysis) on social presence is desirable to ascertain how the development, design, and instruction of online learning moderates the effects of social presence on student outcomes. However, the only attempt to systematically synthesize and quantify the research literature specifically on social presence was a meta-analysis that was limited to studies up until May 2015 (Richardson et al. 2017). Though it provided a measure of the effect social presence has on perceived student learning and satisfaction, the number of studies included was small, and the moderating effects of course design, instructor behaviors, type of online learning environment (e.g., asynchronous, synchronous, blended, or MOOC) were not determined.
As social presence research conducted in a variety of online learning environments continues to increase (Mykota, 2018), the ability to find and generate stable and possibly convergent results for student outcomes of social presence is enhanced.

**Theoretical Framework**

The purpose of the research protocol is to report on a prospective systematic review and meta-analysis that expands on Richardson et al.’s (2017) earlier findings. To accomplish this, a systematic review of the research literature from 1995-2021 will be conducted to examine the relationship between social presence and student outcomes. Data will then be extracted to determine, through a meta-analysis, how the moderating conditions of course design elements, instructor behaviors, and type of online learning environment affect the strength of the relationship. The overarching research questions to be answered are:

1. **How strong is the relationship between social presence and students’ learning in online courses?**
2. **How strong is the relationship between social presence and students’ satisfaction in online courses?**
3. **What moderating conditions (e.g., course design elements, instructor behaviors, and type of online learning environment) affect the strength of these relationships?**

The proposed systematic review and meta-analysis allows for greater predictability in the understanding of the combined effects across studies that social presence has on student outcomes. More importantly, the study is unique in that it will help determine how the aggregated moderating effects of course design, instructor behaviors, and type of online learning environment influence student outcomes due to social presence. In turn, the development, design, and instruction of online learning will be better informed as to which practices facilitate the improvement of social and interpersonal communication necessary to enhance student learning outcomes and create quality online learning environments. Further, the higher education online community
will benefit from a more strategic and informed conceptualization of what constitutes 
social presence. Finally, this research explores how best to interpret the research 
literature on the design of higher education online learning environments, and how to 
facilitate student and instructor engagement in the learning process. This research brief 
reports the research protocol for the proposed systematic review and meta-analysis.

**Objectives**

The objectives of the proposed study are fourfold:

1. Conduct a systematic review of the research on individual studies that statistically 
   report on the relationship between social presence and student learning and student 
   satisfaction outcomes and extract those studies for meta-analysis.
2. Conduct a meta-analysis that statistically summarizes the individual studies 
   extracted from the systematic review that report on the relationship between social 
   presence and student learning and satisfaction outcomes.
3. Determine, as part of the meta-analysis, the effects that course design elements, 
   instructor behaviors, and type of online learning environment have on the 
   relationship between social presence and student learning and student satisfaction 
   outcomes and discover their magnitude of influence as moderators.
4. Support the online higher education community through dissemination of research 
   findings, and implementation of knowledge mobilization activities.

**Methodology**

A systematic review of the literature will be undertaken to obtain an appropriate 
sample of studies from the research for the meta-analysis (i.e., the statistical summary of the 
studies’ results). The sample for the synthesis will be derived from studies that report effect 
sizes for the relationship between social presence and student learning and student 
satisfaction outcomes. The study will occur in two phases: the first phase will involve a
systematic review of the literature while the second phase of the study will comprise the meta-analysis.

**Conducting the Systematic Review**

To conduct the systematic review (Objective 1), the expertise of the research librarian was solicited. Using the research questions outlined earlier, a preliminary list of key concepts will be constructed. The initial key concepts will include terms such as post-secondary, online learning environment, student/teacher affect, and finally social presence. Selected databases will yield literature (i.e., journal articles, conference papers, theses and dissertations) pertinent to a study on social presence and include Web of Science, ERIC Ovid, PsychINFO, and ProQuest Education. Using the list of key concepts, each broad term will be mapped to terms in the first database (ERIC/OVID) to discover specific synonyms to search for in ERIC and subsequent databases. The Boolean operator “and” will be used to build relevant search criteria, and “or” will be used to combine the key concepts to extract relevant articles and documents for import into Endnote 20. Limits will be put on the search to include relevant material from 1995 to December 2021 and key terms will be mapped to be located anywhere in the reference material including the title, abstract, and full text. A manual search of references from past studies relevant to the systematic review and meta-analysis will be conducted along with a Google scholar search to capture any additional records that might be missed. Relevant studies will be exported into Endnote 20 and de-duplicated.

The methods for the systematic review outlined below are modeled after the guidelines provided by The Campbell Collaboration (2020) in the conducting of systematic reviews, the PRISMA checklist for reporting, and the data collection flow diagram for systematic reviews and meta-analysis (Moher, et al. 2009). A three-stage screening process
will be developed for the systematic review to determine the adequacy of studies that might fit the meta-analysis. In the first stage, an initial title, abstract, and keyword screening for social presence studies will focus on student outcomes (e.g., learning and satisfaction), and will be conducted by two reviewers. Discrepancies will be resolved either through consensus or, if needed, involvement of a third reviewer. The second stage of the screening process will involve material previously identified as uncertain. To ascertain if material identified as uncertain from the initial screening is suitable for detailed review will require reading of the full text by the reviewers, with a final consensus reached on their relevancy. The reference lists of studies reviewed in detail will be searched for additional studies to be reviewed that were not included in the database search.

The third stage involves identifying articles suitable for inclusion in the meta-analysis that were reviewed in detail in the first two stages. Inclusion criteria for the meta-analysis is based on the following conditions: the relevancy of the material to social presence and the identified outcomes (student learning and student satisfaction), and the appropriate measurements of the variables reported (e.g., t-values, correlations, regression coefficients, sample size, standard deviations, and means). It is important to have the appropriate measurements of the variables to enable the calculation of the correlation coefficient between social presence and the identified outcomes.

Of those studies screened for inclusion through the above-mentioned process, a full text read will occur, and a data extraction form completed. The form will summarize the study features that include author and year, publication type, participants, course length, discipline area, type of learning environment (asynchronous online, synchronous online, blended, MOOC, etc.), the social presence scale used, sample size, method of data analysis and any potential bias that is reported. A coding scheme will be developed for potential moderator characteristics (e.g., course design elements, instructor behaviors, and type of
online learning environment). To facilitate the analysis, course design elements will be aggregated and coded into categorical variable(s), as will instructor behaviors and the type of online learning environment. By undertaking such a coding scheme, an analysis of categorical moderators is facilitated which helps to explain variation in effect size and differences in results across studies.

The quality of the extracted studies will be determined through a review of their internal, external, and construct validity. For example, are the studies underpinned by a strong social presence theoretical framework? Are the conclusions statistically valid? Are the measures reliable and do they demonstrate strong internal consistency? How are fishing and error rate problems addressed? Is the power of the study provided (i.e., sample size, statistical test, significance, research hypothesis, and effect size)? How are statistical assumptions addressed?

The data extraction form will be piloted by the research team and assessed for completeness, ease of use, and percentage agreement between reviewers targeted at >90%. Based on the pilot testing, any required modifications to the data extraction form will be undertaken to ensure the data necessary to address the research questions is obtained. Data collected in the data extraction process will be housed in a secure data storage environment maintained by the University of Saskatchewan, accessible only to members of the research team. This data management plan ensures the guidelines for the collecting, storing, and sharing of data are in alignment with Social Sciences and Humanities Research Council of Canada’s research guidelines.

**Conducting the Meta-Analysis**

As social presence studies tend to use rating scales as measures of social presence (Biocca, et al., 2003; Chen et al., 2015; Mykota, 2018), for the meta-analysis (Objective 2),
correlational coefficients will be used to calculate the effect size, with transformation of the correlation coefficient to Fisher’s z score to normalize the distribution (Fisher, 1915; Hedges and Olkin, 1985). Besides determination of the effect size, the approximate sampling variance of the effect size is also calculated (see formulas in Borenstein, 2009; Borenstein et al., 2009; 2021; Cheung and Vijayakumar 2016). Two separate meta-analysis will occur, one for student learning outcomes and the other for student satisfaction outcomes. If a study reports both outcomes, they will be treated as independent and entered separately for each outcome. A random-effects model is chosen as the statistical model for analysis of the two outcome measures as it allows effects to differ from study to study with the goal being to estimate the distribution of the effect sizes. In contrast, a fixed effects model assumes that the studies included in the meta-analysis share a common effect size. In this sense, a random effects model for analysis allows studies to have their own population size effects while acknowledging, as in this meta-analysis, that studies using the described search methods are treated as a sample of all the possible studies—other studies might not be identified using the search criteria adopted for inclusion in the meta-analysis (Borenstein et al., 2010; Cheung et al., 2012).

Since a random effects model for analysis is proposed, an a priori test of power will be conducted to determine if the number of studies included are sufficient for statistical power for the testing of the heterogeneity of effect sizes, and average effect sizes (Borenstein et al., 2009; Cheung and Vijayakumar 2016; Valentine et al., 2010). Heterogeneity of effect sizes are determined using a Q test (Cochran, 1954). However, when the number of studies in a meta-analysis is small, the Q statistic has inadequate power to detect heterogeneity. Subsequently, the between groups variance $\tau^2$ statistic will also be calculated. As well, the $I^2$ statistic for determining true heterogeneity, represented by the percentage of total variability of effect sizes due to study differences, will be determined.
(Huedo-Medina et al., 2006). Using these three tests provides a more accurate interpretation of heterogeneity of effect sizes in the meta-analysis sample (Higgins and Thompson, 2002; Higgins et al., 2003).

Next, a moderator analysis (Objective 3) will be undertaken to identify study characteristics that might explain the differences in the relationship between social presence and the two outcomes. Potential characteristics extracted and coded as moderators include course design elements, instructor behaviors, and type of online learning environment. If a single study examines the effect under multiple moderating variables, a separate case will be entered in the data set. To reduce dependence in each analysis, cases for the same level of moderator being examined will be combined as per the guidelines offered by Cooper (2010) and DeCoster (2009). If the moderator categories are unbalanced and the number of studies small (< 20), a pooled estimate, rather than separate estimates, for the residual between studies variance will be used (Rubio-Aparicio et al., 2017). To test the hypothesis that some of the study characteristics act as moderators a mixed-effects analysis, commonly referred to as a meta-regression, will be undertaken. If indeed the regression coefficients are positive, this would suggest that the effect sizes are stronger for studies for which moderator analysis was undertaken. One caveat to this analysis is found in the interpretation of the regression coefficient; it is important the coefficient is not interpreted at the individual level but at the group level (Cheung and Vijayakumar 2016).

A sensitivity analysis will be undertaken to help understand differences between studies and outcomes. Forest plots, for example, will be used to illustrate chronologically how evidence has accumulated over time. Sub-groups of studies can then be analyzed to answer what if questions, by excluding some studies based on a particular criterion of interest. In turn, a visual representation of the data will be presented that could reveal patterns of interest for the sub-group analysis. This might include an examination of earlier
studies on the construct and its effects on student learning and satisfaction with those more recently reported to see how the understanding and measurement of the construct has influenced outcomes over time.

Finally, as it is likely that only studies with significant results are reported and published in the research literature, it is important, especially for the conducting and interpretation of a meta-analysis, to determine if publication bias does exist. To this end, there are several methods for determining publication bias. For purposes of the present meta-analysis, the use of funnel plots, trim and fill procedures, along with fail-safe N enable the ascertainment of the impact that publication bias has on the meta-analysis findings (Rothstein et al., 2006). The software chosen for the analysis is Comprehensive Meta-Analysis (Version 3). Comprehensive Meta-Analysis has been reviewed and is recommended as a software for both the conducting of meta-analysis and for use as a teaching tool with graduate students (Pierce, 2008). Comprehensive Meta-Analysis is one of the few software available that can compute all analysis.

**Limitations**

There are a few limitations that might affect the systematic review and meta-analysis. First, a small sample size could affect the generalizability of results and the ability to determine the heterogeneity of variances for effect sizes. Second, the reporting of study characteristic variables in the extracted records might compromise the ability to conduct a moderator analysis. Third publication bias might be present for studies that report findings related to actual and perceived learning, which could compromise the ability to determine effect sizes.
**Conclusion**

Systematic reviews and meta-analysis are desirable because there is a methodological rigor associated with them. Study protocols are an important aspect of this methodological rigor because they provide clarity surrounding the conducting of a systematic review, inform others as to the research process involved, reduce duplication of research efforts, and lessen potential publication bias. This proposed systematic review and meta-analysis allows for greater predictability in the understanding of the pooled effects across studies that social presence has on student outcomes. In turn, the extracted study designs can be replicated and previous results either confirmed or refuted.

Through knowledge mobilization activities and dissemination of the findings (Objective 4), academic and non-academic audiences involved in higher education online learning will be informed as to how best to design innovative online courses to engage students and instructors in the learning process. This is significant because post-secondary institutions will continue to be better positioned to attract, retain, and graduate students so they can meet the challenges and demands of an evolving society and labor market.

The proposed study will also provide recommendations for rigor in research design so that subsequent research on the evolution and improvement of social presence can be empirically validated. Future comparative research across the social science disciplines that considers learning environment, enrollment, course level, and sub-discipline is also recommended to determine what practices work best in certain situations and what practices can be generalized to all types of online learning environments. By so doing, best practices for quality online learning environments in higher education can be advanced.
References


Moallem, M. (2015, July). The impact of synchronous and asynchronous communication tools on learner self-regulation, social presence, immediacy, intimacy, and satisfaction in collaborative online learning. *The Online Journal of Distance Education and E-Learning, 3*(3), 55–77. [https://tojdel.net/journals/tojdel/articles/v03i03/v03i03-08.pdf](https://tojdel.net/journals/tojdel/articles/v03i03/v03i03-08.pdf)


Author

Dr. David Mykota is a professor at the University of Saskatchewan. Some of his current research interests include online learning program development, program evaluation, social presence, knowledge synthesis, and systematic reviews. He currently instructs undergraduate, certificate, and graduate e-learning courses in special education and educational psychology and is presently engaged in SSHRC-funded research pertaining to the affective domain of e-learning.

Email: david.mykota@usask.ca

Acknowledgements

Funding for this research has been provided by the Social Sciences and Humanities Research Council of Canada’s Insight Grant program.

This work is licensed under a Creative Commons Attribution 3.0 License.