Computers & Education 172 (2021) 104252

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

The influence of technology acceptance, academic self-efficacy, and gender on academic achievement through online tutoring

José Hanham [a,*], Chwee Beng Lee [a], Timothy Teo [b]

[a] School of Education, Western Sydney University, Locked Bag 1797, Penrith, NSW, 2751, Australia
[b] Discipline of Education, College of SHEE, Murdoch University, 90 South Street, Murdoch, WA, 6150, Australia

ARTICLE INFO

Keywords:
Online tutoring
Academic achievement
Self-efficacy
Technology acceptance
Gender

ABSTRACT

The study focused on variables which were posited to capture undergraduate students' experiences with a large-scale online tutoring service, and relationships with the students' perceptions of their academic capabilities and academic performance. A theoretical model incorporating variables from research on *Technology Acceptance* and *Social Cognitive Theory* was developed and tested. A total of 365 undergraduate students from a university located in Sydney, Australia, completed an online survey. Data were analysed using confirmatory factor analysis (CFA) and structural equation modeling (SEM), and multi-group analyses (MGA). The measurement model demonstrated configural, metric and scalar invariance. There were differences between males and females regarding latent means, with females scoring higher than males for facilitating conditions. The regression paths were consistent across males and females (i.e., invariant) in the full structural model. Facilitating conditions was positively associated with the perceived usefulness of technology, which in turn was positively associated with academic self-efficacy. Surprisingly, perceived ease of use did not have a statistically significant association with perceived usefulness. Academic self-efficacy was positively associated with academic achievement. Implications, particularly for online tutoring service providers, are discussed.

1. Introduction

Tutoring has a long history that predates formal classroom teaching and can be traced back at least to ancient Greece (Robinson, Schofield, & Steers-Wentzell, 2005). Historically, tutoring has generally being conceived as a relationship between a person with sufficient knowledge of a subject area (i.e., the tutor) who is able to impart her or his knowledge to improve the content knowledge and/or skills of a less knowledgeable person (i.e., the tutee) in that subject area (Graesser, D'Mello, & Cade, 2011). Although most of the literature on tutoring has focused on one-to-one tutoring between adults and school-age students in synchronous settings, there has been a growing interest in online tutoring over the last several decades (Cao, Yang, Lai, & Wu, 2021; Kulik & Fletcher, 2016; Price, Richardson, & Jelfs, 2007; Richardson, 2009; Steenbergen-Hu & Cooper, 2014; VanLehn, 2011).

There are two general forms of online tutoring. One form is based on advancements in artificial intelligence which led to what was initially described as *Computer Assisted Instruction* (CAI) (Suppes & Morningstar, 1969), and later evolved into what is currently known as *Intelligent Tutoring Systems* (ITS) (Kulik & Fletcher, 2016; Steenbergen-Hu & Cooper, 2014; VanLehn, 2011). This form of online tutoring involves semi-autonomous computer programs providing instructional feedback to tutees. The other form of online tutoring is

* Corresponding author.
E-mail address: j.hanham@westernsydney.edu.au (J. Hanham).

https://doi.org/10.1016/j.compedu.2021.104252
Received 16 July 2020; Received in revised form 4 May 2021; Accepted 28 May 2021
Available online 4 June 2021
0360-1315/© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

based on advancements in communication technologies, which have enabled human tutors to provide instructional feedback to tutees either synchronously or asynchronously via various communication tools (e.g., video conferencing, instant messaging, electronically annotated feedback etc). This form of tutoring has been described as person-to-person online tutoring (Johns & Mills, 2021). Over the last two decades, numerous commercial online tutoring enterprises have emerged offering learners around the globe with opportunities to receive instruction and feedback from human tutors via an online platform (Nelson-Royes, 2015).

Although online tutoring services have proliferated in recent decades, empirical research on person-to-person online tutoring is relatively sparse (Cao et al., 2021; Price et al., 2007; Richardson, 2009). It is currently unclear how tutees' perceptions of the usability of person-to-person online tutoring services are associated with their perceptions for successfully completing academic assignments (i.e., academic self-efficacy) and academic achievement outcomes. This study seeks to address this gap in the research literature.

2. Literature review

Aside from distance education, tutoring has largely been an activity in which human tutors provide instruction to their tutees in synchronous, face-to-face settings. Online tutoring emerged in the 1960s, however, at that time, online tutoring was largely based on specialised computer tutoring programs, rather than humans tutoring tutees via an online system. Online tutoring using human tutors started to appear in the higher education sector in the late 1980s and early 1990s, most notably in the UK Open University in which tutoring was carried out via electronic conferencing and telephone (Mason, 2000). Within a few years of the World Wide Web entering the public sphere in 1993, larger numbers of higher education institutions, as well as private enterprises, for example Net Tutor, established in 1996, became involved in person-to-person online tutoring.

In recent decades there has been a rapid expansion in online tutoring services (Zhang & Bray, 2020). A combination of educational research findings, socio-economic changes, and technological advancements, and most recently a global health pandemic, have played a role in driving this rapid expansion. In terms of research findings, it has been empirically evident since the 1980s (e.g., Cohen, Kulik, & Kulik, 1982) that learners who receive tutoring tend to have higher academic achievement gains than learners who do not receive tutoring. Research in subsequent decades, including several meta-analyses, have revealed that human tutoring in face-to-face settings, as well as online tutoring via intelligent tutoring systems, are associated with greater levels learner achievement than more traditional classroom instruction methods (Kulik & Fletcher, 2016; Steenbergen-Hu & Cooper, 2014; VanLehn, 2011). However, it is important to note that studies involving person-to-person online tutoring have been largely absent from these meta-analyses, in part due there being very few studies on this form tutoring.

Although tutoring appears to promote academic achievement, the mechanisms that explain the association between tutoring and improved academic achievement are not well understood. Whilst direct empirical data may be absent, scholars have inferred that two well-known instructional strategies, feedback and scaffolding, may explain the effectiveness of tutoring on academic achievement (VanLehn, 2011). Feedback is widely recognised as one of the most effective strategies for improving student learning and achievement (Wisniewski, Zierer, & Hattie, 2020). Tutors can provide tutees with feedback about their current levels of performance and/or understanding of content areas. Moreover, tutors can provide tutees with strategies and direction that can improve their performance levels and knowledge states. Feedback is often integrated with scaffolding, which refers to adaptive instructional support (Van de Pol, Volman, & Beishuizen, 2010). Scaffolding is a metaphor adapted from the construction field, which refers to a temporary support structure. In the context of tutoring, a range of instructional scaffolds such as worked examples, alongside other strategies such as questioning, monitoring progress, and motivational support are provided to tutees. Scaffolding seeks to enable tutees to attain learning and achievement gains beyond what they could initially achieve without support (Kleickmann, Tröbst, Jonen, Vehmeyer, & Möller, 2016). Over time, scaffolding is gradually faded so that it is no longer necessary (Van de Pol et al., 2010).

Although a comprehensive understanding of the processes by which tutoring promotes academic achievement has yet to be reached, the fact that there is extensive empirical evidence in support of the academic benefits of tutoring has been important in the promotion, and subsequent growth tutoring services, including online tutoring services. Education is a means for improving social mobility and empirical data shows that increased levels of education correlate with improved socio-economic outcomes, including increased earning potential (Psacharopoulos & Patrinos, 2018). Alongside factors such as socio-economic status, academic achievement is one of the key deciding factors that allow learners to access higher levels of education. Throughout the globe there is an industry known as supplemental education which involves learners accessing additional instruction, primarily through tutoring, to improve their chances of gaining access to higher levels of education (Zhang & Bray, 2018). The tutoring provided through supplemental education generally aligns with the topics covered in the formal curricula of schools and universities, with the main goal of improving learners' academic performance, particularly on high stakes assessments (e.g., university entrance exams). Indeed, research has found that educational aspirations are positively associated with engagement in supplemental education through tutoring (Guill & Lintorf, 2019).

Online tutoring is quickly becoming, if not so already, the preferred mode by which supplemental education is provided to learners (Zhang & Bray, 2020). One reason for this is that online tutoring is generally more affordable than face-to-face tutoring (Takashiro, 2018; Ventura & Jang, 2010) This is likely due in part to fewer overhead costs, as well as the capacity outsource tutoring to countries with lower wages. As examples, *Growing Stars* and *TutorVista* are two companies that have headquarters in the United States and deliver online tutoring to students in the United States, however, their tutors are based in India, which has a highly educated populace, with English as one of their main languages. Differences in the dollar values between India and United States has meant that it is more profitable to run a tutoring business with tutors based in India rather than United States.

In addition to being more affordable, online tutoring is often positioned as being more convenient for both tutors and tutees than face-to-face tutoring (Ventura & Jang, 2010). Convenience is obviously due to the affordances provided by the Internet, but also other technological advancements. Currently, most online tutoring interactions occur through a Learning Management System (Godwin-Jones, 2016), which is software that enables tutors to create and curate content, deliver lessons, assess the work of their tutees and provide feedback on that work. Online tutoring via an LMS is not bound by geographical boundaries and therefore tutees can generally access a tutoring service regardless of where they are located as long as they have access to the Internet. Some prominent LMS providers such as *Canvas* and *Blackboard* possess multiple tools that tutors can use to assist tutees including, learning analytics, plagiarism detection software, accessible content functionality, and online video conferencing. Because online tutoring is delivered through an LMS via the Internet, the hours of operation of online tutoring services are much more flexible, with numerous online tutoring services able to provide on-demand tutoring assistance or at least tutoring feedback within a 24-hr period upon an initial request for tutoring (Nelson-Royes, 2015). Other elements of convenience that have been mentioned include not having to travel to physical premises, as well as safety as a tutor does not have to enter and attend the home of the tutee (Ventura & Jang, 2010).

2.1. Problem statement

Although it is evident that tutoring generally leads to improvement in learner achievement outcomes, most of the research evidence has been obtained from studies of traditional tutoring in face-to-face settings, as well as intelligent tutoring systems (Kulik & Fletcher, 2016; Steenbergen-Hu & Cooper, 2014; VanLehn, 2011). Limited research has been directed towards person-to-person tutoring in online environments (Chappell, Arnold, Nunnery, & Grant, 2015; Johns & Mills, 2021; Price et al., 2007). In the current COVID pandemic, online tutoring has become the most prevalent form tutoring, and in many countries the only means by which learners can access supplemental education (Zhang & Bray, 2020). Whilst online tutoring services may be more affordable, convenient, and in the current climate one the few ways learners can access tutoring, we know little about learners' experiences with person-to-person online tutoring and how they may be associated with their achievement outcomes. A key difference between traditional face-to-face tutoring and person-to-person online tutoring is that interactions between tutors and tutees in the latter are mediated by an online delivery platform. The extent to which the platform is functional and easy to use may play a role in the effectiveness of the tutoring sessions. For example, if tutees find the platform interface difficult to use and navigate, some users may not persist with tutoring. Furthermore, technical issues with the platform prior to, or during a tutorial session, may disrupt learning. From the educational technology literature, factors such as learners' perceptions of the ease of use technology, as well as, the availability of technological support and infrastructure, known as facilitating conditions, have been associated with learners' acceptance and use of technology (Granić & Marangunić, 2019; Song & Kong, 2017). In turn, perceptions of the usefulness of a technology platform have been found to be associated with learners' beliefs in their academic capabilities, that is, their self-efficacy beliefs (Abdullah & Ward, 2016). This is important because there is a long-standing link between self-efficacy beliefs and academic achievement outcomes (Affuso, Bacchini, & Miranda, 2017; Honicke & Broadbent, 2016; Travis, Kaszycki, Geden, & Bunde, 2020). The aim of this study is to investigate how learners' experiences of online tutoring are related to academic outcomes. Specifically, the relationships between facilitating conditions, perceived ease of use, and perceived usefulness with their academic self-efficacy beliefs and academic achievement grades. In addition, we also considered the potential moderating role of gender, which is of interest to researchers given that research has shown that technological acceptance beliefs and self-efficacy may vary according to gender (Huang, 2013; Ong & Lai, 2006; Padilla-Meléndez, Aguila-Obra, & Garrido-Moreno, 2013). Finally, we also considered the relationship between duration, that is how long learners had used the tutoring service and academic achievement.

2.2. Constructs

2.2.1. Perceived usefulness, perceived ease of use, facilitating conditions

To capture learners' perceptions of the usability features of a large-scale online tutoring service, we drew upon two constructs from the *Technology Acceptance Model* (TAM; (Davis, 1989), *perceived usefulness* (PU) and *perceived ease of use* (PEU). Perceived usefulness reflects users' beliefs that technology, for example, specific software, hardware or system, will enhance their productivity. Perceived ease of use refers to users' beliefs about the extent to which a particular form of technology will require effort to learn and operate. We also incorporated the construct, *facilitating conditions*, which originated from Thompson, Higgins, and Howell (1991) and was integrated into the Unified Theory of Acceptance and Use of Technology model (UTAUT; Venkatesh, Morris, Davis & Davis, 2003). This construct refers to factors in the environment, namely, organisational and technological infrastructures that shape users' perceptions of the ease or difficulty of performing tasks (Venkatesh, MorrisDavis, & Davis, 2003). Facilitating conditions include technical support, availability of instructional support, and absence of technical issues (Teo & Noyes, 2014).

2.2.2. Academic self-efficacy

Perceived usefulness, perceived ease of use, and facilitating conditions are variables that have been positioned as predictors of learners' attitudes towards, and intentions to use, technology (Teo & Huang, 2019). Academic performance, rather than attitudes and/or intentions toward technology is a key focus of this study. Consequently, we incorporated the construct, academic self-efficacy (Richardson, Abraham, & Bond, 2012). Self-efficacy is a key construct in Bandura's (1997, 2001) Social-Cognitive Theory and refers to domain and task specific beliefs that people have about their capacity to organise resources and execute courses of action needed to successfully perform tasks. Strong self-efficacy can result in learners behaving in ways that are likely to enhance learning and academic performance. This includes applying extra energy and determination when performing tasks, remaining resolute when confronting obstacles, and choosing difficult over easier learning tasks when given the choice. In contrast, learners with relatively weak self-efficacy are more inclined to behave ways that diminish the likelihood of achieving learning gains and improved academic outcomes. This includes applying minimal effort when performing tasks, opting out of completing tasks when faced with challenges, and preferring easier learning tasks over more challenging learning tasks when given the option. Not only is self-efficacy related to motivational outcomes, it has been shown to be associated with academic achievement (Affuso et al., 2017; Honicke & Broadbent, 2016; Travis et al., 2020), including academic performance in online learning environments (Kitsantas & Chow, 2007; Joo, Lim, & Kim, 2013; see Yokoyama, 2019 for summary).

2.2.3. Academic achievement

Assessment of academic achievement is a core function of higher education institutions around the globe. Academic achievement can be considered a multifaceted construct (Cachia, Lynam, & Stock, 2018) underpinned by a range of variables. These include cognitive factors such as intelligence (Mayes, Calhoun, Bixler, & Zimmerman, 2009) and working memory (Alloway, 2009), non-cognitive factors such as self-efficacy and educational aspirations Lee & Stankov, 2018, as well as by the social-cultural contexts in which students are predominately exposed (Liem, 2019). Academic achievement is measured through assessment tasks, which can take different forms, including exams, essays, reports, class presentations, projects and practicums and so on. The purpose of assessment tasks varies and includes diagnostic assessment which measures pre-existing knowledge, formative assessment which measures student progress, and summative assessment which measures educational attainment (i.e., knowledge and skills) at the end of a unit of study. Summative assessments generally employ an alphanumeric grading/marking scheme as a measure of academic achievement. Although tutors may assist with diagnostic and formative assessment, generally their main purpose is to improve academic achievement outcomes for tutees (Ömeroğulları, Guill & Köller, 2020) on summative tasks, such as high stakes tests (Guill & Lintorf, 2019).

2.2.4. Gender

In this study we were also interested in gender differences, which have been the subject of empirical investigations in both technology acceptance research (e.g., Lin & Yeh, 2019; Ong & Lai, 2006; Padilla-Meléndez, del Aguila-Obra, & Garrido-Moreno, 2013; Sánchez-Franco, 2006), and academic self-efficacy research (e.g., Huang, 2013). In technology acceptance research, Padilla-Meléndez et al. (2013) and Sánchez-Franco (2006) found that relationships between perceived ease of use and perceived usefulness were significantly stronger for males than females. Terzis and Economides (2011) found that females had higher ratings for perceived ease of use and facilitating conditions for a computer based assessment, whereas males had higher ratings for perceived usefulness of a computer based assessment. Similarly, Ong and Lai (2006), found that females gave more weight to perceived of use, whereas, males placed more emphasis on perceived usefulness. Other studies have found minimal or no gender differences regarding technology acceptance variables (Teo, Fan, & Du, 2015, Whitley, 1997).

A meta-analysis of gender differences regarding academic self-efficacy identified a small overall effect indicating that males generally have higher academic self-efficacy than females, though academic self-efficacy varied by academic domains. Females generally have higher self-efficacy for languages, whereas males generally have higher self-efficacy in mathematics, computing and social sciences (Huang, 2013). Given that gender differences have been found for technology acceptance variables and academic self-efficacy, it may reasonably be expected that gender may moderate the relationships between the constructs examined in this study.

2.2.5. Duration of tutoring and academic achievement

One of the earliest meta-analyses on face-to-face tutoring programs (Cohen et al., 1982) as well as more recent meta-analysis on peer assisted tutoring (Leung, 2015), found that the duration of tutoring, that is, how many weeks/months a learner participated in tutoring program, was a factor in their academic achievement outcomes. Interestingly, research (Leung, 2015) suggests that participation in tutoring programs of shorter duration (i.e., less than 10 weeks), had stronger effects related to academic achievement than tutoring programs of longer duration (i.e., greater than 10 weeks).

2.3. Description of the tutoring service

The online tutoring service examined in this study provides tutoring specifically for academic writing assignments (i.e., essays, reports). To access the tutoring service, students are provided a link on a Blackboard Learning Management System in which they can submit a written assignment anytime (24/7). During the submission process, students are provided with a range of questions including free-response and drop-down menu options. As a first step, students are asked to provide a general description of the assignment task that they are required to complete. Following this, the students are asked how far they have progressed with the assignment and are given three options to choose: “It’s my first draft”, “Nearly there” and “Almost ready to hand in”. Students are then asked to identify the specific subject that the assignment is located, followed by identifying the type of assignment (e.g., essay, report, case study etc). Students are then given four options for which they can receive tutoring assistance for their written assignment: structure, choice of language, spelling/grammar and referencing. Students can receive tutoring assistance with one or more of these options. As a final question students are asked: Is there anything else you think we should know? Once this information is submitted, the written assignment is then allocated to a trained tutor who will provide written feedback generally within a 24-hr period on one or more of the aforementioned options. After receiving the feedback, students are then asked to complete a short evaluation survey that includes indicating their level of satisfaction with the service (1 extremely satisfied to 5 extremely unsatisfied) and whether or not they believed that they received the help that the needed by clicking on one of the following icons:

3. Formulation of hypotheses and research questions

In this study, the usability of the tutoring service is captured using two constructs from the technology acceptance literature, perceived ease of use (PEOU) and facilitating conditions (FAC). Items regarding PEOU were direct statements regarding the ease of use of the online tutoring service (e.g., "The online tutoring service is easy to use"). Facilitating conditions contained items concerning the infrastructure aspects that contribute to the ease of use of the tutoring service ("*There is always* support *when I need help using online tutoring service*"). Previous research has found that both perceived ease of use (Abdullah & Ward, 2016) and facilitating conditions (Teo, 2019) were positively associated with users' perceptions of the usefulness of technology. Based on these findings it posited here that the greater the perceived usability of the of online tutoring service, the more likely that learners will perceive the online tutoring service to be useful. The following two hypotheses reflect this position:

Hypothesis 1. Perceived ease of use will be positively associated with perceived usefulness

Hypothesis 2. Facilitating conditions will be positively associated with perceived usefulness

Tutoring can assist students in a number of ways. In this study, we were interested in learners' perceptions about the extent to which the online tutoring service helped them structure and complete their assignments more effectively, as well as, illuminating their understanding of the requirements of the assignment and improving their conceptual knowledge. These were the key elements of the perceived usefulness of the online tutoring service, which contribute to improvements in various aspects of academic literacy. Past research in technology acceptance has found that perceived usefulness is positively correlated with learners' self-efficacy beliefs (Chen, Chang, Chen, Huang, & Chen, 2012). In light of this, it is posited that the more useful the online tutoring service is perceived to be in helping students successfully complete their assignments, the more likely that learners will perceive themselves capable of attaining higher achievement and learning outcomes. This is reflected in the following hypothesis:

Hypothesis 3. Perceived usefulness will be positively associated with academic self-efficacy

The research literature shows that academic self-efficacy is related to achievement and performance outcomes in educational settings, including online learning environments (Joo et al., 2013; Kitsantas & Chow, 2007). As such, the following hypothesis is advanced:

Hypothesis 4. Academic self-efficacy will be positively associated with academic achievement.

The research hypotheses are diagramed in the theoretical model presented in Fig. 1.

3.1. Research questions

The main research question of this study is concerned with the theoretical model (see Fig. 1).

RQ1. To what extent to does the data support the hypothesised relationships described in the theoretical model?

There are several questions related to gender. The findings regarding gender differences in technology acceptance research and self-efficacy research are mixed (Padilla-Meléndez et al., 2013; Sánchez-Franco, 2006; Whitely, 1997) and thus rather than positing hypotheses, it was considered more appropriate to put forward research questions. The second research question focuses on measurement, and concerns the comparability of responses between males and females for items designed to measure technology acceptance variables and academic self-efficacy.

RQ2. To what extent are the users' responses to measure their experiences of an online tutoring service and academic self-efficacy, invariant across males and females?

The third research question is concerned with potential gender differences in the average scores for each of the latent constructs.

RQ3. What are the gender differences in average scores (latent means) for each of the latent constructs?

The fourth research question is concerned with the potential moderating role of gender on the relationships between the variables

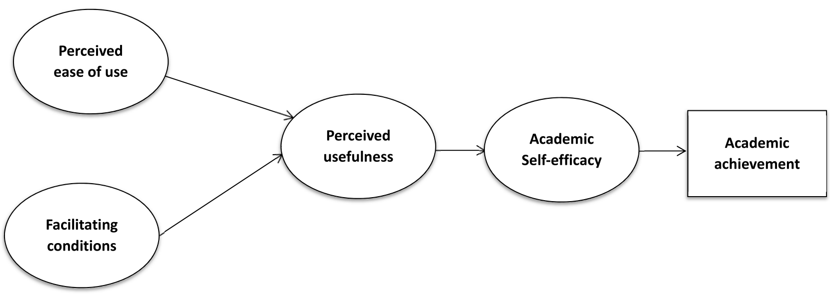

Fig. 1. Theoretical model.

in the theoretical model.

RQ4. How does gender moderate the relationships between the variables specified in the theoretical model?

The final research question relates to the duration of tutoring. Studies of traditional tutoring programs suggest that the duration of tutoring is related to academic achievement (Cohen et al., 1982; Leung, 2015). However, it does not appear that the relationship between how long learners have used an online tutoring service and their academic achievement outcomes has not been investigated.

RQ5. What is the relationship between the duration of participation in an online tutoring service and academic achievement?

4. Research methodology

A cross-sectional design was employed utilising data from an online survey. The hypotheses and research questions 1 and 5 were tested using structural equation modeling (SEM). SEM was used as it has several advantages over more traditional multivariate techniques such as regression. This includes the fact that SEM accounts for measurement error and enables researchers to evaluate the degree of fit between the theoretical model and sample data, which is generally not as feasible with more traditional multivariate techniques (Hair, Black, Babin, & Anderson, 2019). Although SEM is not used to examine casual relationships in this study, relatively close fit between the sample data and theoretical data does enhance claims regarding potential casual relationships (Bollen & Pearl, 2013). Multi-groups analysis (Muthén & Muthén, 1998-2017) is considered to be an appropriate method for testing for measurement invariance and moderating effects and was used to address research questions, 2, 3, and 4.

4.1. Participants

A total of 365 undergraduate students from a large, metropolitan university located in Sydney, Australia, participated in the study. All of the participants were current users of an online tutoring service, offered by an external provider, which was accessed through the university's student support service. The sample comprised 72.6% females (n = 265) and 27.4% males (n = 100). Participants ranged in age from 18 to above 65 years, with the largest number of participants 66.3% (n = 242) in the age range of 18–24 years. Participants were widely spread across eight organisational schools in the university (Arts and Social Sciences = 38.1%, Business Studies = 9.15%, Computing, Engineering, Mathematics = 2.29%, Education = 13.31%, Law = 2.29%, Nursing = 17.05%, Psychology = 5.41%, Sciences = 11.23, Miscellaneous e.g., research methods = 1.02%).

4.2. Procedure

Protocols for the study were approved by the Human Research Ethics Committee of the university. Email invitations with an information package regarding the study were sent to users of a third-party online tutoring service that had an official partnership with the student support services at the university. All participants were given a $10 (AUD) iTunes voucher for taking part in the survey. Out of approximately 500 potential responses, we had full data from 365 respondents (73% response rate).

4.3. Instrument

All data for this study were collected using a questionnaire administered through the online application, *Survey Monkey*. The questionnaire comprised three sections. The first section captured demographic information, that is, age, gender, subject, length of use (i.e., duration) of the online tutoring service, as well as a measure of academic achievement which was the participant's self-reported grade for the assignment for which they received tutoring. The second section contained 13 items presented on a Likert-type scale, ranging from "1" (*strongly disagree*) to "7" (*strongly disagree*). The items used to measure perceived usefulness (n = 6 items) and perceived ease of use (n = 3 items) were adapted from Davis (1989). Four items were used to measure facilitating conditions and were adapted from Thompson et al. (1991). The third section consisted of 10 academic self-efficacy items several of which were adapted from the *MSLQ* (see Pintrich, Smith, Garcia, & McKeachie, 1991) and several items were developed specifically for this study. The participants were asked to respond to the self-efficacy items based on their experiences with the online tutoring service. Bandura's guidelines (2006), were measured on an 11-point percentage scale ranging from 0% (*not at all confident*) to 100% (*completely confident*).

The four theoretical constructs of interest were operationalized in the following ways: i) *perceived usefulness*, reflected the extent to which the online tutoring service was perceived as useful in helping students to understand, structure, and complete their assignments on time; ii) *ease of use*, captured students' perceptions of the usability the online tutoring service; iii) *facilitating conditions*, *facilitating conditions*, also captured students' perceptions of the usability online tutoring service, but with reference to infrastructure support, including technical support and instructional support; iv) academic *self-efficacy*, measured students' perceived confidence that they could perform a range of academic tasks in their subject area of study based on their experiences with the online tutoring service. The questionnaire items are presented in Appendix 1.

4.4. Data analysis

Data analyses were preformed using IBM SPSS version 26 (IBM Corp, 2019) the open source statistical program R, version 3.6.3 (R Core Team, 2017), and Mplus version 8.3 (Muthén & Muthén, 1998-2017). The default procedure in Mplus, full information maximum likelihood (FIML) was used to handle missing data. Descriptive analyses were conducted for the demographic variables, specifically the participants' age and gender. For testing of measurement and structural models, the robust maximum likelihood (MLR) estimator (Satorra & Bentler, 2001) was used as it is robust with regard to violations of normality. To assess model fit, several measures of fit were consulted, specifically the comparative fit index (CFI) and the Tucker Lewis Index (TLI), the root mean square of error approximation (RMSEA) and the standardised root mean square residual (SRMR). CFI and TLI values range from 0 to 1, with values above 0.90 and 0.95 are suggestive of acceptable to excellent fit (e.g., Hu & Bentler, 1999). For RMSEA and SRMR values that are closer to zero suggest better fit, with values less than 0.05 for RMSEA indicating good fit, with values up to .08 generally considered to indicate acceptable fit, and values over 0.10 indicating poor fit (MacCallum, Browne, & Sugawara, 1996). SRMR values of 0.08 and below are considered good fit.

Measurement invariance tests were carried out to address research question 2. Furthermore, measurement invariance tests served a methodological purpose related to research questions 3 and 4. As a precondition for testing for differences in latent means (RQ3.) and the potential moderating influence of gender regarding the structural paths (RQ4.), measurement variance up to scalar invariance must first be established (Muthén & Muthén, 1998-2017). Measurement invariance tests were carried using a hierarchical sequence, commencing with configural invariance tests followed by metric and scalar invariance tests. Following scalar invariance tests, latent means difference tests were then conducted. Because the sample sizes were unbalanced between males and females, we adopted a subsampling procedure advocated by Yoon and Lai (2018), which adjusts for unbalanced designs for measurement invariance analyses. This procedure involves using the R software program to generate 100 subsamples from the larger group, which in this study is the female cohort. Each of the 100 subsamples is the same size as the male sample. Configural, metric and scalar tests were run with all of the subsamples. The Mplus program calculates the averages of the fit statistics across samples to generate means for chi-square, RMSEA and CFI.

To test whether a particular model is invariant, one may run a chi-square difference test (Wang & Wang, 2012), with a non-significant $\Delta\chi2$ statistic indicative of invariance across groups. However, the $\Delta\chi2$ statistic is generally oversensitive to small model differences. Consequently, an adjunct set of measures of fit, namely CFI and RMSEA have been identified collectively as a viable alternative for assessing invariance (Parada, 2019). When comparing a restrictive model with the preceding model, a decrease of 0.01 or less for CFI and an increase in RMSEA of 0.015 or less is generally considered to provide support for invariance across groups (Parada, 2019). The chi-square statistic is reported in our description of each model (configural, metric, scalar), though we assess invariance specifically through examining the degree of change in RMSEA and CFI.

For the substantive purpose of addressing research question 4, tests for the potential moderating effects of gender were examined. This involved testing a baseline model of the full structural model in which all parameters were freely estimated. This model was then compared to a model in which the all regression paths in the structural model were constrained to be equal across males and females. To assess potential differences between the models, changes in RMSEA and CFI were again consulted.

5. Results

5.1. Descriptive analysis

In terms of duration, descriptive analyses revealed that the majority of the participants (64%) had used the online tutoring service for at least one semester, with 28.5% having used it for one semester, 20.8% for two semesters and 15.1% using the service for more than a year. For the factors in the overall sample, the skew indexes ranged from −1.146 to −0.560 and kurtosis indexes ranged from −0.075 to 1.672. These values fall within recommended (Kline, 2009) guidelines for univariate normality. The means and standard deviations for the factors for the full sample, as well as the female and male samples are reported in Table 1.

As seen in Table 1, the means for the technology acceptance variables (response range 1–7) are all above the mid-point (>3.5), for the overall, male and female samples. Similarly, the mean for self-efficacy (response range 0%–100%) is also above the mid-point for the overall, male and female samples.

5.2. Test of the measurement model

The fit indices for the confirmatory factor analysis of the full measurement model were acceptable with relatively good fit $\chi2 = 531.855$ (df = 224, p < .01), RMSEA = 0.06 (90% CI: 0.06–0.07), CFI = 0.95, TLI = 0.95, and SRMR = 0.03..

Table 1
Means and standard deviation of the latent variables.

Latent variable	Overall sample	Female sample	Male sample
Perceived Usefulness	M = 4.98 *SD* = 1.44	M = 5.05 *SD* = 1.39	M = 4.80 *SD* = 1.57
Perceived Ease of Use	M = 5.62 *SD* = 1.43	M = 5.72 *SD* = 1.43	M = 5.36 *SD* = 1.41
Facilitating Conditions	M = 5.12 *SD* = 1.43	M = 5.23 *SD* = 1.48	M = 4.82 *SD* = 1.59
Academic Self-Efficacy	M = 8.15 *SD* = 1.81	M = 8.01 *SD* = 1.76	M = 8.34 *SD* = 1.94

* Perceived Usefulness, Perceived Ease of Use, Facilitating Conditions were measured on Likert scale 1–7.
Academic self-efficacy measured on percentage scale 0%–100%.

5.3. Convergent validity

As seen in Table 2, the factor loadings for the items in the measurement model were well above minimum accepted thresholds (Hair et al., 2019), ranging 8.34 to 9.44 (mean = 0.90; median = 0.91). The Average Variance Extracted (AVE) was used to test for convergent validity and reliability was assessed using the Composite Reliability (CR) measure. As a rule of thumb, AVE and CR that are equal to or greater than 0.50 are considered adequate (Teo & Noyes, 2014). As reported in Table 2, all AVE and CR scores were greater than the minimum accepted threshold values.

5.4. Discriminant validity

The correlations between the latent variables are presented in Table 3. The correlations ranged from 0.38 to 0.86, with the average correlation of 0.60 The square roots of the AVEs were calculated to assist with assessing discriminant validity. Using the guidelines suggested by Fornell, Tellis, and Zinkhan (1982), values in the diagonal elements of the correlation matrix (i.e., the square root of the AVE for a given construct) are greater than the off-diagonal elements in the corresponding rows and columns. An examination of the data in Table 3 provides support for discriminant validity for all of the latent variables.

5.5. Test of the structural model and hypotheses

Structural equation modeling was used to test the hypothesised relationships described in our initial model (see Fig. 1.). The fit indices for the structural model suggested relatively good fit $\chi 2 = 605.992$ (df = 270, $p < .01$), RMSEA = 0.06 (90% CI: 0.05–0.07), CFI = 0.95, TLI = 0.95, and SRMR = 0.04. Table 4 reports the standardised path coefficients and hypotheses tests for the theoretical model. Three out of the four hypotheses were supported by the data. The path from perceived ease of use to perceived usefulness was not statistically significant, therefore Hypothesis 1 was not supported. The path from facilitating conditions to perceived usefulness was statistically significant, and based on Cohen's (1992) guidelines, large in magnitude. Thus, Hypothesis 2 was supported. The path between perceived usefulness and self-efficacy was statistically significant and moderate in magnitude, thus providing support for Hypothesis 3. The path between self-efficacy and student's grades was statistically significant, moderate in magnitude, and thus providing support for Hypothesis 4. Finally, in order to address research question 5, we explored the relationship between the length of participation in the tutoring service (i.e., duration) and learners' grades. A small, but statistically relationship between the two variables (path coefficient = .10, z-value = 1.96) was found, indicating that the longer the participants had used the tutoring service, the higher their grades. The R square values indicated that self-efficacy explained 20% of the variance in academic achievement. Perceived

Table 2
Standardised factor loadings, AVE, CR.

Factors	Items	Standardised Loadings
Perceived Usefulness (AVE = .77; CR = .95)		
	PU1	.898
	PU2	.915
	PU3	.874
	PU4	.844
	PU5	.879
	PU6	.845
Perceived Ease of Use (AVE = .85; CR = .95)		
	PEOU1	.915
	PEOU2	.912
	PEOU3	.944
Facilitating Conditions (AVE = .78; CR = .93)		
	FAC1	.919
	FAC2	.924
	FAC3	.834
	FAC4	.844
Academic Self-Efficacy (AVE = .74; CR = .98)		
	SE1	.854
	SE2	.893
	SE3	.901
	SE4	.911
	SE5	.917
	SE6	.924
	SE7	.931
	SE8	.916
	SE9	.913
	SE10	.892

Table 3
Correlations matrix and square roots of AVEs.

CFA Variable	1	2	3	4
1. Perceived Usefulness	(.87)			
2. Perceived Ease of Use	.72	(.92)		
3. Facilitating Conditions	.79	.86	(.88)	
4. Academic Self-Efficacy	.44	.39	.38	(.86)

*The square roots of the AVEs are in parentheses.

Table 4
Standardised regression path coefficients and hypotheses results.

Hypothesis	Path	Path Coefficient	z value	Result
H1	PEOU → PU	.16	1.60	Not supported
H2	FAC → PU	.66	6.57*	Supported
H3	PU → SE	.44	7.98*	Supported
H4	ASE → Grade	.43	7.99*	Supported

N.B. PU = Perceived Usefulness, PEOU = Perceived Ease of Use, FAC = Facilitating.
Conditions, SE = Academic Self-efficacy. * denotes statistically significant p-values <.05.

usefulness explained 19% of the variance in self-efficacy. Perceived ease of use and facilitating conditions explained 64% of the variance in perceived usefulness.

5.6. *Test of measurement invariance*

Based on the guidelines from Brown (2006), prior to carrying out an any invariance tests, the initial measurement model was fitted to two separate datasets, one comprised of the male participants only, and the other, of the female participants only. The fit statistics for male and female only models were acceptable (Male: $\chi2 = 367.842$ (df = 224, $p < .01$), RMSEA = 0.08 (90% CI: 0.07–0.10), CFI = 0.93, TLI = 0.92, and SRMR = 0.04. Female: $\chi2 = 550.731$ (df = 248, $p < .01$), RMSEA = 0.08 (90% CI: 0.07–0.08), CFI = 0.94, TLI = 0.93, and SRMR = 0.04). Following this, a subsampling procedure (see Yoon & Lai, 2018) was applied to address the uneven sample distribution between males and females. As discussed previously, this involved generating 100 subsamples of the female cohort using the open source R program, with each of the 100 subsamples being the same size as the male sample.

The first model tested was the baseline model, as known as the configural model. The factor loadings, intercepts and error variances are freely estimated in the configural model. This model is the least restrictive model and informs us whether or not the same items measure the same construct for males and females. Following this, metric invariance was tested through adding a restriction to the model, in which factor loadings are constrained to be equal across groups. The metric model informs us whether or not the factor loadings are equivalent across males and females. A scalar model was then tested which adds another restriction wherein the equivalency of intercepts between groups is constrained to be equal. As we used a subsampling procedure, the MPlus program calculates the means for the fit statistics for across the 100 subsamples. The means for Chi-square, RMEA and CFI are reported in Table 5, and as can be seen that there no changes to RMSEA and CFI between the models, indicating measurement invariance for gender.

5.7. *Tests for latent mean differences*

To address RQ2, a test of latent mean differences was conducted. A precondition for testing for differences in latent means, the intercepts of the reflective variables (i.e, the items) should be invariant groups (i.e., the scalar model). As noted by Teo et al. (2015), "the purpose of achieving over identification of the factors (necessary condition for testing model fit), analysis of latent mean differences requires that the factor intercepts for one group be fixed to zero" (p. 244). In this study, the male cohort served as the reference group, and therefore, the factor means for this cohort were constrained to zero. The results of the latent mean difference tests are reported in Table 6. There were statistically significant differences one latent variable, with females having higher mean values for facilitating conditions.

Table 5
Fit statistics of models used in measurement invariance tests.

Model	$_{Mean\ \chi}2a$	df^b	Mean RMSEA	Mean CFI
Configural	813.794	448	.09	.92
Metric	836.421	467	.09	.92
Scalar	856.348	486	.09	.92

Table 6
Tests for averaged latent means differences.

Factor	Difference estimate	Z-value
Perceived Usefulness	.289	1.40
Perceived Ease of Use	.369	1.87
Facilitating Conditions	.470	2.06*
Academic Self-Efficacy	-.181	−0.74

* denotes statistically significant p-values <.05.

5.8. Tests for moderating effects of gender in the structural model

The steps to test whether gender moderates the structural relationships (i.e., regression path coefficients), is analogous to the steps that were used for testing measurement invariance. As a first step, a baseline model is specified in which all the parameters are freely estimated. The statistics of this model are compared with a model in which all of the regression path coefficients are constrained to be equal. If there is significant variance between the two models as indicated by a decrease of more than .01 for CFI and an increase of more than 0.015 in RMSEA, then a series of comparative model tests is applied. This will involve constraining each path in a model one at a time and comparing the fit of those models with the baseline model. Table 7 shows the means for the following fit statistics for both models: chi-square, CFI and RMSEA. Examination of Table 7 shows that there are no changes in mean scores for RMSEA and CFI between the baseline and constrained model, and as such, the structural model was considered invariant. That is, the associations between the latent variables are consistent across females and males.

6. Discussion

This study investigated how university students' experiences with a large online tutoring service were associated with their academic self-efficacy and academic achievement. Three key aspects of user experience were measured: i) direct measures of *perceived ease of use* of the online tutoring service; ii) *facilitating conditions*, namely infrastructure support that contributed to the ease of use of the online tutoring service; iii) the *perceived usefulness* of the tutoring service in assisting students to understand the requirements of the assignment and helping them complete their assignments in a timely and effective manner. The main research question (RQ1) was concerned with to what extent the data supported the hypothesised relationships depicted in the theoretical model. Three out of four hypotheses were supported, suggesting that to a large extent, the data provide good support for the theoretical model. Notably, academic self-efficacy explained nearly one-fifth (R square = 19%) of the variance in academic achievement, which the R square value approaching what is considered what moderate effect (Ferguson, 2009). This result is generally consistent the extant research literature (see Honicke & Broadbent, 2016) with academic self-efficacy generally been found to have moderate association with academic achievement across a range of subject-content areas and educational settings. Perceived usefulness also accounted for nearly one-fifth (R square = 19%) of the variance in academic self-efficacy. Perceived ease of use and facilitating conditions accounted for nearly two-thirds (R square = 64%) of the variance in perceived usefulness, which is considered a strong effect in social science research (Ferguson, 2009).

Discussing each respective hypothesis, no support was found for Hypothesis 1, which predicted that the perceived ease of use of the online tutoring service would be positively associated with the perceived usefulness of the online tutoring service. Support was found for Hypothesis 2, with a statistically significant path from facilitating conditions to perceived usefulness. The fact that facilitating conditions, but not perceived ease of use, was positively associated with student's perceptions of the usefulness of the tutoring service is noteworthy. In general, in technology acceptance models, perceived of ease of use is posited to be associated with perceived usefulness (Abdullah & Ward, 2016). All the participants had some prior experience with online tutoring service, with the majority of the participants (64%) having already used the service for one semester prior to data collection. As such, it is likely that most of the participants found the online tutoring service easy to use, and therefore infrastructure aspects (i.e., facilitating conditions) were more salient features relevant to the perceived usefulness of the online tutoring service.

Support was found for Hypothesis 3, in that perceived usefulness of the online tutoring service was positively associated with academic self-efficacy. Various aspects of the online tutoring service were measured to capture its perceived usefulness. This included its role in helping students structure and understand the requirements of their assignments, its role in improving student learning and conceptual understanding, and its role in helping students complete their assignments on time. Essentially, the online tutoring service contributed to these aforementioned aspects through instructional feedback to students. Feedback fits into the realm of verbal persuasion, which is a key source of self-efficacy (Chan & Lam, 2010). It is important to note that the four sources of self-efficacy articulated by Bandura (1997), mastery experiences, vicarious experiences, verbal persuasion, and physiological and emotional

Table 7
Fit statistics of models used in structural invariance tests.

Model	Mean χ^2	*df*	Mean RMSEA	Mean CFI
Baseline SEM	894.860	496	.09	.92
Regression paths constrained SEM	900.323	500	.09	.92

state are not always distinct sources, but can also influence each other (Usher & Pajares, 2008). For example, it is possible that verbal feedback from more expert/knowledgeable others can contribute to learners' perceptions of mastery. During tutoring sessions, feedback from tutors may inform tutees of their progress towards learning a concept (conceptual knowledge) or performing a procedure (procedural knowledge) (Narciss et al., 2014). This feedback from the tutors can be used by the tutees as evidence of their progression towards mastery of a concept or procedure. It is important to acknowledge that in this study, we did not capture that nature of feedback provided by the tutees. Future research, including qualitative studies, may explore the nature of feedback provided by tutors in online environments and its impact on the relationship between the perceived usefulness of the tutoring service and learners' self-efficacy beliefs.

The data from the study provided for Hypothesis 4 which posited that academic self-efficacy would be positively associated with academic achievement, with students' self-reported grades used as the measurement for academic achievement. In this study, the more the students perceived themselves capable of understanding, analysing, elaborating and applying concepts taught in their unit of study, the more likely they were to score a higher grade for their assignments. The positive association between academic self-efficacy and academic achievement is well-established (Affuso et al., 2017; Honicke & Broadbent, 2016; Travis et al., 2020). There is also growing evidence of the positive association between academic self-efficacy and academic achievement in online settings (Joo et al., 2013; Kitsantas & Chow, 2007; Yokoyama, 2019). It is likely that the tutoring received by many of the participants contributed to advancing their capabilities for understanding, analysing, elaborating and applying their conceptual knowledge. These cognitive skills align with Bloom's taxonomy of learning objectives and are key components of academic writing assignments (Olena, 2017). It reasonable to expect that the more capable the tutees in this study perceived themselves in these cognitive skills, the more likely they were to produce written assignments of high quality.

Three research questions addressed in this study were related to gender. RQ2 was concerned with to what extent participant responses to the online survey were comparable across males and females. In terms of the consistency responses to survey, the series of progressively stringent invariance tests indicated that the male and female participant cohorts had comparable results regarding the factor structure (configural invariance). It was also found that each latent variable influenced their reflective indicator (item) to a similar extent across both males and females (metric invariance), and furthermore, that the item intercepts were similar across the male and female cohorts (scalar invariance).

RQ3 was concerned question with potential gender differences in the average scores for each of the latent constructs. Establishing scalar invariance is an important precondition for exploring latent mean differences (Van de Schoot, Lugtig, & Hox, 2012). Based on latent means comparisons, there was a statistically significant gender difference for facilitating conditions with females having statistically higher ratings for this latent variable than males. In other words, females were more likely than males to perceive that there was appropriate infrastructure support to assist them when needed. Likewise, the reverse is true in that males were less likely than females to perceive that infrastructure and support was available when needed. The higher ratings by females for facilitating conditions in online environments has been found in previous research (Terzis & Economides, 2011).

RQ4 was concerned with whether or not gender moderated the relationships between the variables in the theoretical model. The multi-group analyses did not find that gender had any significant influence on the relationships between the variables in the theoretical model, thus answering the fourth research question.

RQ5 was concerned with the relationship between the duration of participation in the online tutoring service and academic achievement. The results of the structural equation modeling indicated that there was a small, but still statistically significant relationship between duration in the online tutoring service and academic achievement. This result aligns with findings from studies of traditional tutoring programs (e.g., Cohen, et al., 1982).

6.1. Implications for theory and practice

There are a number of implications for theory and practice that stem from the results of this study. Currently, theoretical understanding of the variables that are associated with learners' experiences with online tutoring and their academic achievement outcomes is limited. The findings from this study suggest that theorisations around students' experiences with online tutoring may benefit from incorporating variables from the technology acceptance literature, particularly, those that take into account infrastructure aspects of online tutoring services. The results from the study also to add to the research literature (Richardson et al., 2012) attesting to the important association between academic self-efficacy and academic achievement.

From a practical viewpoint, when educational institutions are in the process of selecting an external online tutoring service provider to include as part of a suite of measures to support students, a key consideration should be the quality of infrastructure support that comes with tutoring service. This includes the extent to which instructional and technical support is readily available to respond to student needs. Likewise, this should be a key consideration of potential and existing online tutoring service providers, especially given that universities around the globe are seeking to provide students with anywhere, anytime access to educational support. Given the gender differences regarding facilitating conditions in this study, online tutoring services with large numbers of female tutees should ensure that careful attention is placed on infrastructure support for tutees.

Another practical implication concerns the relationship between the perceived usefulness of the online tutoring service and students' academic self-efficacy beliefs. The literature suggests that self-efficacy is associated with positive academic outcomes in online environments (Bradley, Browne, & Kelley, 2017). The fact that the more students perceived the online tutoring service useful in helping them to understand, structure and complete their assignments, the stronger the beliefs in their academic capabilities, is important because most higher education students must now complete some or all of their studies online. Indeed, given the current Covid-19 pandemic, online learning is the only avenue for which many universities can deliver, and students access, learning content.

Not all students are equipped with the necessary knowledge, skills and resources to successfully manage online learning (Bol & Garner, 2011). Online tutoring services, if structured and delivered effectively, are likely to be invaluable to universities in helping them support many students who have to deal with a rapid transition from predominately face-to-face modes of instruction to fully online modes of instruction (Zhang & Bray, 2020).

6.2. Limitations and directions for future research

There are a number of limitations associated with this study that should be acknowledged. First, all the of the participants were from a single university. To increase the generalizability of the findings, future research should draw from a larger number of randomly selected universities. Second, only a limited number of variables were included in the study. There may be other important variables that impact students' academic self-efficacy and academic achievement. Past academic performance has been shown to predict self-efficacy and academic achievement (Hwang, Choi, Lee, Culver & Hutchinson, 2016). It might well be that the effects of academic self-efficacy might diminish significantly when past academic performance is included in the analysis. General cognitive ability is another variable that has been shown to be related to both academic self-efficacy and academic achievement (Pajares & Kranzler, 1995). Interestingly, Pajares and Kranzler (1995) found that self-efficacy was still significantly associated with academic performance after controlling for general cognitive ability. Third, the data are cross-sectional and therefore causality cannot be inferred. Indeed, some of the variables included in this study, such as self-efficacy beliefs are fluid and change over time (Hanham, McCormick, & Hendry, 2020). Accordingly, future studies should employ longitudinal designs to capture potential changes in self-efficacy and other variables over time. Fourth, the data for the study were largely based on self-reports. Future research may incorporate observational and/or interview methods to capture more in-depth insights into the different aspects of perceived ease of use, facilitating conditions, and the perceived usefulness of online tutoring services. Fifth, academic achievement was assessed based on participants' self-reported grades. Although research has found that self-reported grades positively correlate with actual grades (Sticca et al., 2017) and further that there was no incentive for the participants to misreport their grades as the survey was anonymous, it has been found that lower-performing learners may be less likely to accurately report their grades (Kuncel, Crede'& Thomas, 2005).

7. Conclusions

This study sought to contribute to the literature concerning person-to-person online learning and learner achievement outcomes. The findings from the study suggest that infrastructure aspects of an online tutoring service were a salient factor in students' perceptions of the perceived usefulness of the online tutoring service. In turn, perceptions of the usefulness of the online tutoring service were associated with students' perceptions of their academic capabilities (i.e., self-efficacy), which in turn, were associated with their academic achievement outcomes. This study builds on a small, but necessary body of research literature on variables that influence student experiences and academic achievement with online tutoring.

Credit author statement

José Hanham: Conceptualization, Methodology, Investigation, Data Analyses, Review & Editing, Writing of original draft. Chwee Beng Lee: Conceptualization, Methodology, Investigation, Review & Editing. Timothy Teo: Writing - review & editing.

Declaration of competing interest

All authors declare that they have no conflict of interest.

Acknowledgements

This work was supported by YourTutor (now known as Studiosity) (Grant Number: REDI ID: P00024466).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compedu.2021.104252.

References

Abdullah, F., & Ward, R. (2016). Developing a general extended technology acceptance Model for E-Learning (GETAMEL) by analysing commonly used external factors. *Computers in Human Behavior, 56*, 238–256. https://doi.org/10.1016/j.chb.2015.11.036

Affuso, G., Bacchini, D., & Miranda, M. C. (2017). The contribution of school-related parental monitoring, self-determination, and self-efficacy to academic achievement. *The Journal of Educational Research, 110*, 565–574. https://doi.org/10.1080/00220671.2016.1149795

Alloway, T. P. (2009). Working memory, but not IQ, predicts subsequent learning in children with learning difficulties. *European Journal of Psychological Assessment, 25*, 92–98. https://doi.org/10.1027/1015-5759.25.2.92

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: Freeman.

Bandura, A. (2001). Social-cognitive theory: An agentic perspective. *Annual Review of Psychology, 52*, 1–26.
Bandura, A. (2006). Guide to constructing self-efficacy scales. In F. Pajares, & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307–337). Greenwich, CT: Information Age.
Bol, L., & Garner, J. K. (2011). Challenges in supporting self-regulation in distance education environments. *Journal of Computing in Higher Education, 23*, 104–123. https://doi.org/10.1007/s12528-011-9046-7
Bollen, K., & Pearl, J. (2013). Eight myths about causality and structural equation models. In S. L (Ed.), *Handbook of causal analysis for social research*. Netherlands: Springer.
Bradley, R. L., Browne, B. L., & Kelley, H. M. (2017). Examining the influence of self-efficacy and self-regulation in online learning. *College Student Journal, 51*, 518–530.
Brown, T. (2006). *Confirmatory factor analysis for applied research*. NY: The Guilford Press.
Cachia, M., Lynam, S., & Stock, R. (2018). Academic success: Is it just about the grades? *Higher Education Pedagogies, 3*, 434–439. https://doi.org/10.1080/23752696.2018.1462096
Cao, J., Yang, T., Lai, I. K.-W., & Wu, J. (2021). *Student acceptance of intelligent tutoring systems during COVID-19: The effect of political influence*. The International Journal of Electrical Engineering & Education. https://doi.org/10.1177/00207209211003270
Chan, J., & Lam, S.-F. (2010). Effects of different evaluative feedback on students' self-efficacy in learning. *Instructional Science, 38*, 37–58. https://doi.org/10.1007/s11251-008-9077-2
Chappell, S., Arnold, P., Nunnery, J., & Grant, M. (2015). An examination of an online tutoring program's impact on low-achieving middle school students' mathematics achievement. *Online Learning, 19*, 37–53.
Chen, M. Y., Chang, F. M. T., Chen, C. C., Huang, M. J., & Chen, J. W. (2012). Why do individuals use e-Portfolios. *Educational Technology & Society, 15*, 114e125.
Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155–159. https://doi.org/10.1037/0033-2909.112.1.155
Cohen, P. A., Kulik, J. A., & Kulik, C.-L. C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal, 19*, 237–248. https://doi.org/10.3102/00028312019002237
Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*, 319–340.
Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice, 40*, 532–538. https://doi.org/10.1037/a0015808
Fornell, C., Tellis, G. J., & Zinkhan, G. M. (1982). Validity assessment: A structural equations approach using partial least squares. *Proceedings, American Marketing Association Educators' Conference*.
Godwin-Jones, R. (2016). Looking back and ahead: 20 years of technologies for language learning. *Language, Learning and Technology, 20*, 5–12. Retrieved from http://llt.msu.edu/issues/june2016/emerging.pdf.
Graesser, A. C., D'Mello, S., & Cade, W. (2011). Instruction based on tutoring. In R. E. Mayer, & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 408–426). New York: Routledge Press.
Granić, A., & Marangunić, N. (2019). Technology acceptance model in educational context: A systematic literature review. *British Journal of Educational Technology, 50*(5), 2572–2593. https://doi.org/10.1111/bjet.12864
Guill, K., & Lintorf, K. (2019). Private tutoring when stakes are high: Insights from the transition from primary to secondary school in Germany. *International Journal of Educational Development, 65*, 172–182. https://doi.org/10.1016/j.ijedudev.2018.08.001
Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th). Hampshire, UK: Cengage Learning.
Hanham, J., McCormick, J., & Hendry, A. (2020). Project-based learning groups of friends and acquaintances: The role of efficacy beliefs. *The Journal of Educational Research, 113*, 133–144. https://doi.org/10.1080/00220671.2020.1756729
Honicke, T., & Broadbent, J. (2016). The influence of academic self-efficacy on academic performance: A systematic review. *Educational Research Review, 17*, 63–84. https://doi.org/10.1016/j.edurev.2015.11.002
Huang, C. (2013). Gender differences in academic self-efficacy: A meta-analysis. *European Journal of Psychology of Education, 28*, 1–35. https://doi.org/10.1007/s10212-011-0097-y
Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55. https://doi.org/10.1080/10705519909540118
Hwang, M. H., Choi, H. C., Lee, A., Culver, J. D., & Hutchison, B. (2016). The relationship between self-efficacy and academic achievement: A 5-year panel analysis. *Asia-Pacific Education Researcher, 25*, 89–98. https://doi.org/10.1007/s40299-015-0236-3
Johns, C., & Mills, M. (2021). Online mathematics tutoring during the COVID-19 pandemic: Recommendations for best practices. *Primus, 31*, 99–117. https://doi.org/10.1080/10511970.2020.1818336
Joo, Y. J., Lim, K. Y., & Kim, J. (2013). Locus of control, self-efficacy, and task value as predictors of learning outcome in an online university context. *Computers in Education, 62*, 149–158. https://doi.org/10.1016/j.compedu.2012.10.027
Kitsantas, A., & Chow, A. (2007). College students' perceived threat and preference for seeking help in traditional, distributed and distance learning environments. *Computers in Education, 48*, 383–395. https://doi.org/10.1016/j.compedu.2005.01.008
Kleickmann, T., Tröbst, S., Jonen, A., Vehmeyer, J., & Möller, K. (2016). The effects of expert scaffolding in elementary science professional development on teachers' beliefs and motivations, instructional practices, and student achievement. *Journal of Educational Psychology, 108*, 21–42. https://doi.org/10.1037/edu0000041
Kline, R. B. (2009). *Becoming a behavioral science researcher. A guide to producing research that matters*. New York, NY: Guilford Press.
Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of intelligent tutoring systems: A meta-analytic review. *Review of Educational Research, 86*, 42–78. https://doi.org/10.3102/0034654315581420
Kuncel, N. R., Credé, M., & Thomas, L. L. (2005). The validity of self-reported grade point averages, class ranks, and test scores: A meta-analysis and review of the literature. *Review of Educational Research, 75*, 63–82. https://doi.org/10.3102/00346543075001063
Lee, J., & Stankov, L. (2018). Non-cognitive predictors of academic achievement: Evidence from TIMSS and PISA. *Learning and Individual Differences, 65*, 50–64. https://doi.org/10.1016/j.lindif.2018.05.009
Leung, K. C. (2015). Preliminary empirical model of crucial determinants of best practice for peer tutoring on academic achievement. *Journal of Educational Psychology, 107*, 558–579. https://doi.org/10.1037/a0037698
Liem, G. A. D. (2019). Academic performance and assessment. *Educational Psychology, 39*, 705–708. https://doi.org/10.1080/01443410.2019.1625522
Lin, P.-H., & Yeh, S.-C. (2019). How motion-control influences a vr-supported technology for mental rotation learning: From the perspectives of playfulness, gender difference and technology acceptance model. *International Journal of Human-Computer Interaction, 35*, 1736–1746. https://doi.org/10.1080/10447318.2019.1571784
MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130–149.
Mason, R. (2000). From distance education to online education. *The Internet and Higher Education, 3*, 63–74. https://doi.org/10.1016/S1096-7516(00)00033-6
Mayes, S. D., Calhoun, S. L., Bixler, E. O., & Zimmerman, D. N. (2009). IQ and neuropsychological predictors of academic achievement. *Learning and Individual Differences, 19*, 238–241. https://doi.org/10.1016/j.lindif.2008.09.001
Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus user's guide* (8th ed.) Los Angeles, CA: Muthén & Muthén.
Narciss, S., Sosnovsky, S., Schnaubert, L., Andrès, E., Eichelmann, A., Goguadze, G., et al. (2014). Exploring feedback and student characteristics relevant for personalizing feedback strategies. *Computers & Education, 71*, 56–76. https://doi-org.ezproxy.uws.edu.au/10.1016/j.compedu.2013.09.011.
Nelson-Royes, A. M. (2015). *Why tutoring?: A way to achieve success in school*. Lanham, Maryland: Rowman & Littlefield.
Olena, K. (2017). Bloom's taxonomy and its role in academic writing and reading skills training at English classes. *Science & Education, 22*, 111–116. https://doi.org/10.24195/2414-4665-2017-2-18

Ong, C.-S., & Lai, J.-Y. (2006). Gender differences in perceptions and relationships among dominants of e-learning acceptance. *Computers in Human Behavior, 22*, 816–829. https://doi.org/10.1016/j.chb.2004.03.006

Padilla-Meléndez, A., del Aguila-Obra, A. R., & Garrido-Moreno, A. (2013). Perceived playfulness, gender differences and technology acceptance model in a blended learning scenario. *Computers & Education, 63*, 306–317. https://doi.org/10.1016/j.compedu.2012.12.014

Pajares, F., & Kranzler, J. (1995). Self-efficacy beliefs and general mental ability in mathematical problem-solving. *Contemporary Educational Psychology, 20*, 426–443. https://doi.org/10.1006/ceps.1995.1029

Parada, R. H. (2019). Assessing perceived school support, rule acceptance and attachment: Evaluation of the psychometric properties of the School Belonging Scale (SBS). *Educational and Child Psychology, 36*, 106–115.

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ)*. Ann Arbor: Michigan: National Centre for Research to Improve Postsecondary Teaching and Learning.

Price, L., Richardson, J., & Jelfs, A. (2007). Face-to-face versus online tutoring support in distance education. *Studies in Higher Education, 32*. https://doi.org/10.1080/03075070601004366

Psacharopoulos, G., & Patrinos, H. A. (2018). Returns to investment in education: A decennial review of the global literature. *Education Economics, 26*, 445–458. https://doi.org/10.1080/09645292.2018.1484426

R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Richardson, J. T. E. (2009). Face-to-Face versus online tutoring support in humanities courses in distance education. *Arts and Humanities in Higher Education, 8*(1), 69–85. https://doi.org/10.1177/1474022208098303

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*(2), 353–387. https://doi.org/10.1037/a0026838

Robinson, D. R., Schofield, J. W., & Steers-Wentzell, K. L. (2005). Peer and cross-age tutoring in math: Outcomes and their design implications. *Educational Psychology Review, 17*, 327–362. https://doi.org/10.1007/s10648-005-8137-2

Sánchez-Franco, M. J. (2006). Exploring the influence of gender on the web usage via partial least squares. *Behaviour & Information Technology, 25*, 19–36. https://doi.org/10.1080/01449290500124536

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66*, 507–514. https://doi.org/10.1007/BF02296192

Song, Y., & Kong, S.-C. (2017). Investigating students' acceptance of a statistics learning platform using technology acceptance model. *Journal of Educational Computing Research, 55*, 865–897. https://doi.org/10.1177/0735633116688320

Steenbergen-Hu, S., & Cooper, H. (2014). A meta-analysis of the effectiveness of intelligent tutoring systems on college students' academic learning. *Journal of Educational Psychology, 106*, 331–347. https://doi.org/10.1037/a0034752

Sticca, F., Goetz, T., Bieg, M., Hall, N. C., Eberle, F., & Haag, L. (2017). Examining the accuracy of students' self-reported academic grades from a correlational and a discrepancy perspective: Evidence from a longitudinal study. *PloS One, 12*. https://doi.org/10.1371/journal.pone.0187367. e0187367.

Suppes, P., & Morningstar, M. (1969). Computer-assisted instruction. *Science, 166*, 343–350. doi: 10.1126/science. 166.3903.343.

Takashiro, N. (2018). Technology use and middle school students' participation in shadow education in Japan. *International Psychology Bulletin, 22*(4), 14–21.

Teo, T. (2019). Students and Teachers' Intention to Use Technology: Assessing Their Measurement Equivalence and Structural Invariance. *Journal of Educational Computing Research, 57*, 201–225. https://doi.org/10.1177/0735633117749430

Teo, T., Fan, X., & Du, J. (2015). Technology acceptance among pre-service teachers: Does gender matter? *Australasian Journal of Educational Technology, 31*, 235–251. https://doi.org/10.14742/ajet.1672

Teo, T., & Huang, F. (2019). Investigating the influence of individually espoused cultural values on teachers' intentions to use educational technologies in Chinese universities. *Interactive Learning Environments, 27*, 813–829. https://doi.org/10.1080/10494820.2018.1489856

Teo, T., & Noyes, J. (2014). Explaining the intention to use technology among pre-service teachers: A multi-group analysis of the unified theory of acceptance and use of technology. *Interactive Learning Environments, 22*, 51–66. https://doi.org/10.1080/10494820.2011.641674

Terzis, V., & Economides, A. (2011). The acceptance and use of computer based assessment. *Computers in Education, 56*, 1032–1044. https://doi.org/10.1016/j.compedu.2010.11.017

Thompson, R. L., Higgins, C. A., & Howell, J. M. (1991). Personal computing: Toward a conceptual model of utilization. *MIS Quarterly, 15*, 125–143. https://doi.org/10.2307/249443

Travis, J., Kaszycki, A., Geden, M., & Bunde, J. (2020). April 2). Some stress is good stress: The challenge-hindrance framework, academic self-efficacy, and academic outcomes. *Journal of educational Psychology*. Advance online publication. https://doi.org/10.1037/edu0000478

Usher, E. L., & Pajares, F. (2008). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of Educational Research, 78*, 751–796. https://doi.org/10.3102/0034654308321456

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist, 46*, 197–221. https://doi.org/10.1080/00461520.2011.611369

Van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher–student interaction: A decade of research. *Educational Psychology Review, 22*, 271–296. https://doi.org/10.1007/s10648-010-9127-6

Van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology, 9*, 486–492. https://doi.org/10.1080/17405629.2012.686740

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. G. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*, 425–478. https://doi.org/10.2307/30036540

Ventura, A., & Jang, S. (2010). Private tutoring through the internet: Globalization and *offshoring. Asia Pacific Education Review, 11*, 59–68. https://doi.org/10.1007/s12564-009-9065-5

Wang, J., & Wang, X. (2012). *Structural equation modeling applications using Mplus*. Chichester: John Wiley & Sons Ltd.

Whitley, B. E., Jr. (1997). Gender differences in computer-related attitudes and behavior: A meta-analysis. *Computers in Human Behavior, 13*, 1–22.

Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology, 10*, 3087. https://doi.org/10.3389/fpsyg.2019.03087

Yokoyama, S. (2019). Academic self-efficacy and academic performance in online learning: A mini review. *Frontiers in Psychology, 9*, 2794. https://doi.org/10.3389/fpsyg.2018.02794

Yoon, M., & Lai, M. H. C. (2018). Testing factorial invariance with unbalanced samples. *Structural Equation Modeling: A Multidisciplinary Journal, 25*, 201–213. https://doi.org/10.1080/10705511.2017.1387859

Zhang, W., & Bray, M. (2018). Equalising schooling, unequalising private supplementary tutoring: Access and tracking through shadow education in China. *Oxford Review of Education, 44*, 221–238. https://doi.org/10.1080/03054985.2017.1389710

Zhang, W., & Bray, M. (2020). Comparative research on shadow education Achievements, challenges, and the agenda ahead. *European Journal of Education, 55*. https://doi.org/10.1111/ejed.12413