Online tutoring has been shown to benefit student achievement, academic independence, and reduction of anxiety (Mendoza & Kerl, 2021; Hrastinski et al., 2019). Ample studies support tutoring for enhancing student success. Few studies, however, address the evaluation of effective tutoring, with an objective to prevent ineffective tutoring. The aim of this research was to build and test a meta-model that identified the hallmarks of one-on-one high-quality online tutoring. Using an exploratory-confirmatory mixed method-ground theory (MM-GT) design, a theoretical model was constructed from implicit and explicit models found in current research. A second model, formatively and traditionally built as a grounded theory model from generous data sharing by Tutor.com of archived session transcripts, was used to construct a meta-model for testing and data. Following use of confirmatory factor analysis and a mixed methods experiment, three latent variables: tutor efficacy, student centered, and instructional were identified. The items for evaluation, the hallmarks of high-quality online tutoring were identified as: (a) Tutor efficacy of platform technology and tools; (b) Use of multiple modalities when applicable; (c) Engagement evidenced as multiple and appropriately paced interactions; (d) Inquiry and incorporation of student's prior knowledge; (e) Support as encouragement of student efforts; and (f) Ensuring understanding through clarifying and checking-in with student. From the Tutor.com, implications for tutor screening and training are evident for those tasked with overseeing tutoring programs. Further research beyond higher education, comparing subjects and exploring frequencies using the 5E model are suggested.

Introduction

Educational psychology Benjamin Bloom once claimed that students who receive one-on-one tutoring reach achievements several standard deviations above their classmates (Bloom, 1984). Since then, online tutoring has been proven to benefit students at all levels and is considered “the most effective education intervention ever to be subjected to rigorous evaluation” (Kraft & Falken, 2021). While there are a fast-growing number of online tutoring services due to an influx of educational funding, there are a few companies that have been serving learners and working with tutors for decades.

Online tutoring studies have revealed multiple benefits, demonstrating positive impacts on student achievement, self-efficacy, and course retention (Mendoza & Kerl, 2021). Tutoring has been perceived by learners as a positive experience. Students feel their coursework improves, report an increased likelihood of course engagement, which leads to higher levels of confidence in the material. These indicators support program retention. In addition, learners reported reduced anxiety and the gain of more transferable skills from participating in one-on-one tutoring. The evidence for individual tutoring is considered convincing (Hrastinski et al., 2019).
An acceptable understanding is that a tutor is someone who can provide, “personal, social, academic, and psychological support . . . promoting positive student outcomes” (Munley, Garvey, & Mcconnell, 2010 as cited by Mendoza & Kerl, 2021). The reality, however, is that not all tutors can or do, and not all sessions build academic independence.

**Problem**

With the online tutoring industry nearing a $50 billion dollar-a-year business in the United States, tutoring programs are inherently unequal (Kraft & Falken, 2021). While ample studies support the benefits of effective tutoring, few studies address the impact of ineffective tutoring. Discussions are rarely found about inadequate tutoring that impedes learning, nor on the data and generalizations found that further frustrated learners who had reached out for tutoring services, only to be disappointed by the process and outcome.

Since tutors are not held to the same hiring practices nor evaluation and accountability as professors and classroom teachers, there are questions about the management of tutor quality. For those that oversee programs, evidence-based models are needed to ensure high quality tutoring for all students in all sessions.

**Purpose**

The purpose of this exploratory-confirmatory-mixed methods-grounded theory (MM-GT) research study was to build and test a model that best identified the hallmarks of high-quality one-on-one online tutoring. This involved building a theoretical model from implicit and explicit models about online tutoring from research, building a formative and traditionally grounded theory model, and combining these into a meta-modelling integration (Shim et al., 2021). Finally, through use of confirmatory factor analysis and a mixed methods experiment, a final and robust meta-model for assessing tutor quality was developed. Our objective was to produce this much-needed model, useful in evaluating online tutoring sessions and perhaps, the screening, training, and evaluating of tutors, so every student can benefit from effective tutoring.

**Research Question**

The primary question that guided this research study was: What theoretical model, grounded in qualitative and quantitative data, reveal hallmarks of one-on-one high quality online tutoring? During the three phases of this research design: exploratory, confirmatory, and model testing, more research questions were asked and answered.

**Design**

The exploratory-confirmatory mixed method-ground theory (MM-GT) design is appropriate for the building of theories. This design begins with an inductively informed literature-based theoretical meta-model. Unlike traditional ground theory designs, the MM-GT method allows for
a priori discovery (Shim et al., 2021), given the impossibility of not having prior knowledge, and for these researchers, who have spent years evaluating tutoring sessions, we could not suggest otherwise.

Next, we constructed a traditional grounded theory model independently of the first model and from actual tutoring session transcripts. These two models were then integrated into a next phase meta-model which was then tested using a mixed-methods experiment (MM-EXP). From these findings, a final integrated model of the hallmarks of high-quality tutoring was produced.

**Phase 1: Exploratory**

Two independent and qualitative exploratory strands from two separate sets of data were collected and analyzed. One meta-model was constructed using published research (from peer-reviewed journal articles) and employed CMap Tools concept map making software. Another model, using a more traditional grounded theory approach, was constructed using the Relationship of Inquiry framework for guidance and engaged archived transcripts for open and axial coding. Through comparison and integration of the two, a final phase 1 meta-model was constructed. This would be tested in phase 3.

**The literature**

Online tutoring is a dynamic and complex phenomenon. Viewpoints can include those of the student, the tutor, and those tasked with ensuring sessions are educationally safe and sound. There are studies that have analyzed reflections from student viewpoints, analyzed reflections from tutor viewpoints, and from administrators that oversee tutoring programs. For this study, we chose to approach the examination of tutoring session transcripts beginning with unpacking of the TPACK framework; the technological, content, and pedagogical knowledge associated with online learning experiences. This framework, developed by Mishra & Koehler (2006), is used to train, and evaluate pre-service and in-service classroom teachers. When the TPACK model is understood and incorporated, educators are said to be successfully integrating all necessary considerations to ensure the most educationally successful online interactions (Rakes et Al., 2022). We began with these lenses.

**Technological Knowledge**

Sembiring’s model of effective online tutoring found that learners perceive a tutor’s learning strategy as influenced by their perception of the technology, rational for using Internet, and perception of media support (2018). Technology does indeed provide the setting for online tutoring. A tutor’s familiarity within that setting is their technological knowledge.

Heinrich et al. (2014) found the setting for a tutoring program impacted effectiveness. The setting should be easily accessible and navigable by an array of learner styles and preferences. Provisions and capabilities of tools can include an area to communicate, via chat, using text, emojis, audio, and/or video. These platforms commonly include a board to work on, a tool bar with choices for writing, tools like rulers or graphing capabilities, and the capability to share
information such as images, files, or screen shares. For consideration is how valuable the tutor’s familiarity with these tools is. Is knowing how to transition between capabilities and tools, incorporating multiple modalities into the tutoring session, and overcoming any difficulty the learner may be having with tools and/or setting, valuable to learners in high quality online tutoring?

**Content Knowledge**

Content knowledge, both conceptual and procedural, is what learners seek when requesting tutoring support. More specifically, learners seek accuracy followed by gainfulness, traceability, and promptness of information (Sembiring, 2018). Tutors, therefore, are perceived as on-demand, credible, content matter experts in specific subjects.

Bledsoe et. al. (2021) found that students studying different disciplines valued instructor effectiveness differently. In humanities-based tutoring, for example, learners valued relationships and teaching methods between instructor and student while in the tutoring of sciences, factual knowledge and performance were more valued. For consideration is how valuable the tutor’s content and procedural knowledge is. Is accepting student requests when the question/problem is revealed, using accurate vocabulary, rules, formulae, and references/resources, and assisting over transferring to another tutor, valuable to learners in high quality online tutoring?

**Pedagogical Knowledge**

Tutors, like traditional classroom teachers, use pedagogic strategies that are critical for learning, including accessing prior knowledge and stimulating student thinking (Hrastinski et al., 2018). Strategies that enhance learning, self-efficacy, and student success all incorporate active learning (Bransford, Brown, & Cocking, 2000). Online tutoring sessions are inherently active as students bring tasks and engage in dialogue or discourse with tutors (Jansson et al., 2021). Active learning involves constructive cognition (Piaget’s constructivism) for the rearrangement, assimilation, and changes in an individual’s mental schema. This is impacted by the social interactions and language used between learner and tutor as a shared learning experience (Vygotsky’s social constructivism). Interchanges between the tutor and learner are as important to gaining new understanding or knowledge as well as the knowledge itself. This implies, “The journey is just as important as the destination” (Lynch, 2016).

Attributes that have been identified by learners in successful tutoring included organization, engagement/enthusiasm/interest/expression, rapport/interaction/concern, respect of students, interaction/questioning, workload/difficult and learning/perceived outcome (Bledsoe et al., 2014). Competencies found to overlap with master teacher effectiveness from students’ viewpoints are problem-solving, availability, communication, creativity, individual consideration, social awareness, feedback, professionalism, and conscientiousness (Catano & Harvey, 2011). For consideration is how valuable the tutor’s pedagogical knowledge is. Is organization, engagement, interaction, availability, communication, professionalism, and conscientiousness valuable to learners in high quality online tutoring?
**Inquiry-Based**

Engaging students in the learning process can be accomplished using an inquiry-based approach. Duran & Duran’s (2004) 5 E model for inquiry-based learning involved using questions asking students to engage, explore, explain, elaborate, and evaluate. Engaging active learning with a task and accessing prior knowledge, exploring the phenomenon, and creating ideas, explaining the new knowledge, elaborating about understanding and applications, and evaluating what has been learned through reflection (Duran & Duran, 2004). For consideration is how valuable inquiry is. Is the questioning of students to engage, explore, explain, elaborate, and evaluate valuable to learners in high quality online tutoring?

**Questions**

A critical part of teachers’ work is asking questions (Jansson et al., 2021). According to these authors, individual tutoring leads to students who outperform traditionally taught students because learners are exposed to better questions by tutors who focus on reasoning, deeper understanding, and eliciting responses from the individual they are working with. A Wilcoxon Signed-Ranks test \( Z = -5.80, p < 0.01 \) indicated that the tutors were more likely to ask questions \( M = 6.2, \ SD = 1.7, \ N = 363 \) as compared with lecturing \( M = 5.7, \ SD = 2.0, \ N = 371 \) (Hrastinski et al., 2021). From tutors’ viewpoints, questions are necessary to encourage students to reflect, understand the level of knowledge of the student, understand what mistakes students have made, understand different areas and connections of the subject area, and guide students toward problem-solving (Hrastinski, Cleveland-Innes, & Stenbom, 2018). Chen et al. (2019) found that tutors asked 4-5 times more questions than students in tutoring sessions. They also found that the more questions asked, the more likely it was a successful tutoring dialogue. For tutoring sessions in which tutors lecture, answer-give, or are overly active at instructing student interactions, student interactions were reduced (Mazzolini & Maddison, 2007). Others found some answer-giving is useful for enriching facts (Stenborn, et al., 2016). Feng et al. (2016) recommended facilitation of discourse be emphasized in the mid-phase, while direct instruction scaffolding is needed in the last phase. Is inquiry-based questioning over answer-giving valuable to learners in high quality online tutoring?

The type of question a tutor asks is also important. Using closed questions is ideal for rote memorization but does not employ higher order thinking like open questions (Dohrn & Dohn, 2018). Open-ended questions, regardless of being direct or indirect, were found to be more likely to elicit student responses (Jansson et al., 2021). For example, these authors note that when the question, “How can I help you?” was present, students felt that the tutors were more likely to engage them during the process \( U = 6994, \ p = 0.02 \). Statistical significance was also found with the question type, “Do you have an idea how to solve?” \( U = 3050, \ p = 0.03 \). These findings were based on exploring question types on conversation intensity, approach to tutoring, perceived satisfaction and perceived learning in online tutoring using archived data from a math coaching project (Jansson et al., 2021).

**Interactions**
Successful tutoring dialogues have also been measured using an indication of student satisfaction, often accomplished using a rating system (Chen et al., 2019). These researchers used 325 features of input, based on 15,000 sessions, to predict tutoring success. These researchers found that positive sentiment through utterances by the tutor made a dialogue more likely to be successful and lead to a higher student satisfaction (Chen et al., 2019). More specifically, the dialogue features included: efforts invested by tutors/students, informativeness of tutor/student utterances, readability level of tutor utterances, tutor responsiveness, questions asked by tutors/students, entrainment level of a tutorial dialogue, the positive sentiment level of tutor/student utterances, and student experiences in using the tutoring service (Chen et al., 2019). However, these authors believed more research was needed to verify which of these factors really impacted student satisfaction. Chappell et al. (2015) claimed more was needed to overcome the challenges of how to engage student interaction in online synchronous tuition. Janson et al. (2021) echoed this concern stating there was limited understanding in what characterized the inquiry process in online tutoring sessions since most research was based on outcome measures and perceptions. For consideration is how valuable a tutor’s ability to create an engaging and satisfying interaction is. Is encouragement and positive utterances valuable to learners in high quality online tutoring?

**Timeliness**

Successful dialogues between tutors and learners were highly associated with sessions in which students did not wait long to be assisted, nor one in which students waited to receive responses from tutors while in session. Tutor responses were an average of five seconds faster in successful dialogues than in failed dialogues (Chen et al., 2019). These same authors observed that successful dialogues were 50% longer than the failed dialogues, indicating longer engagement and higher satisfaction. Others have also indicated that the length of the tutoring session was fundamental and contributed to the effectiveness of online tutoring (Heinrich et al., 2014). As Kraft & Falken (2021) stated, sustained relationships between tutors and students are key. For consideration is how valuable timeliness and utterances are. Are utterances, length of session, pacing in session, length between responses valuable to learners in high quality online tutoring?

**Findings:** A literature-driven meta-model

Using CMap Tools software from the Florida Institute for Human & Machine Cognition (IHMC), models, concepts, ideas, and recommendations from the literature were iteratively reworked using concept mapping to produce the following model. See Figure 1.

**Figure 1**

*Literature-driven model*
The Grounded Theory Model

For this step, we sought transcripts from previous tutoring sessions that had been archived. The Princeton Review’s Tutor.com provided these valuable resources. As an early pioneer of the online tutoring industry, Tutor.com has been at the forefront of online tutoring for over 20 years. To date, there have been more than 20 million archived session transcripts involving multiple grade-level subjects, including many Humanities and STEM subjects, academic as well as technical and available in multiple languages. In the last few years, tutor pools have averaged 3000 - 3500 individuals at a given time. Weekly sessions average around 30,509.

Archived session transcripts were provided, stripped of all identifiers save the first name of learner and first name of tutor. The data pulled represented 5 years, from 2014 to 2019. From this data, we filtered for sessions rated “5” on a scale from 1 to 5, given by users when asked to rate their satisfaction with the session completed. While Tutor.com serves corporate, governmental, libraries, and K-12 institutes, only transcripts for sessions conducted for higher educational students were utilized.

Of these, we selected session requests for the subject, Algebra \((n=7598)\). We filtered for highly rated “5” sessions that had both a tutor comment left after the session as well as a student comment left after the session to access qualitative feedback from the session, given after it ended. This left us with 1137 sessions. Since transcripts contained a chat log and images from a board, we chose to filter out audio sessions with the reasoning that audio recordings would bring up considerations involving voice, intonation, sound, background, etc. Very few (less than 10%) were audio-based. This may be an area for further study. Finally, we felt session transcripts that were under 10 minutes and those over 50 minutes may either not yield enough points of data, or too many. It was noted that 36% of these best sessions were over an hour
while 3% were under ten minutes. This filter yielded 736 transcripts. Using a randomizing formula in Microsoft Excel, 30 random sessions were generated.

Using Microsoft Excel, we recorded and organized definitions, descriptions, codes, constructs, including reflections, questions, and insight from the process. Through open coding, followed by axial coding, this iterative theoretical sampling process led to multiple categories that we believed contributed to highly valued online tutoring sessions. When we felt the data was saturated and categories explicit, the process was stopped. The emergent categories were as follows.

Findings: A grounded theory-model

(a) Communication

Style—Communication style differed by tutor. Some tutors were very precise with their written language and complete sentences while others used a more abbreviated style and even emojis, however these were used sparingly. In reflection, the theme professional and appropriate kept returning, as in maintaining appropriate language skills and not entertaining off-topic discussions, even when the learner may not have used complete sentences, etc.

Interaction – In many transcripts, there was a demonstration of flow; a momentum of back-and-forth communication that was readily occurring between participants. This discussion flow consisted of moving forward through problem-solving steps. Many of the tutors’ responses were interpreted as positive, encouraging, and supportive, regardless of if the learner demonstrated an incorrect answer, a lack of confidence, and/or doubt.

(b) Content In all of the sessions reviewed, tutors demonstrated both willingness and ability to accept session requests that contained topics from the basic concepts listed in their subject area.

Tutor self-efficacy –Tutors in this sample seemed at ease and demonstrated a familiarity with conceptual and procedural knowledge. In none of the sessions did the tutor claim they could not help, decline to address the problem, avoid learners’ questions, nor pass the session to another tutor. In a vast majority of sessions, the tutor seemed confident and sounded like a subject matter expert.

Learner styles – All of the sessions included a transcript of a chat log. In all of the session transcripts reviewed, the board was used to visually “work” the problem and/or provide an image or resource, like a formula for quadratic equations, for example. In almost every case, writing tools were used over typed or text responses on the board.

(c) Approach The sub-categories below represent the evidence observed from an examination of interactions between tutor and learner that yielded productive, successful solutions.

Assessing learner and task– A consistent theme included tutors identifying needs of learners by asking questions as simple as, “How can I help you”? “What have you done so far”, “What have you learned in class about solving this problem”¿. While
infrequent, there were instances of tutors repeating what the student said they needed, reminiscent of stating objectives prior to the tutoring session.  

**Scaffolding** – The most successful of these high-rated sessions as evidenced by the students’ explicit gratitude and appreciation, tended to have more back-and-forth questions and answers between tutor and student. Tutors seem to guide learners using information provided bit-by-bit conceptually or procedurally as opposed to more verbose and long chunks of information.  

**Clarifying** – Similarly, tutors’ explanations appeared clear and seemingly supported since students rarely challenged the tutor for follow-up clarity on any explanation. The only follow-up by the learner seemed to occur when the tutor asked a question the learner could not readily answer and was looking for more direction by the tutor. Tutors in these sessions asked for clarity in the areas of the problem being presented, to reveal what they knew about the problem, or to find out if the learner understood what they just did. Tutors did this with what seemed like a series of check-ins.  

Returning to the literature, as expected in traditional grounded theory constructivism (Shim et al., 2021), we found support of model categories as well as further insight. From this, we incorporated an understanding of online tutoring using the ROI framework. The Relationship of Inquiry (ROI) framework (Stenborn et al., 2016), an adaptation of the Community of Inquiry framework (Garrison, Anderson & Archer, 2000) is a method to describe meaningful online learning experiences. The four dimensions of the ROI framework are teaching presence, cognitive presence, social presence, and emotional presence (Stenborn et al., 2016). Cognitive presence, as one element, is represented by four sub-categories: triggering event, exploration, integration, and resolution. These were further broken down into indicators such as triggering events which may be stating a problem or changing direction. Another, exploration, was considered exchanging information, integration as connecting ideas and performing computations, and resolution as achieving a solution. The element, teaching presence, was broken in three categories: design and organization (establishing interactions, setting parameters), facilitating discourse (stimulating constructive inquiry, and assessing process), and direct instruction (providing steps, summarizing the discussion). Social presence was based on two categories: open communication (acknowledging, trivial expressions, greetings and building links) and relationship cohesion (greetings, vocatives and building links). Finally, emotional presence, based on three categories: activity emotion (building links, emotions about inquiry), outcome emotion (regarding outcome consequence) and directed effectiveness (?) (emotions towards the other person) (Stenborn et al., 2016).  

For a final series of activities in phase one, the two models were examined in reference to each other and a new meta-model, based on capturing categories from individual nodes of the concept map, emergent categories from coding actual transcripts, and further understanding from the literature was constructed. The focus was on designing a model that would capture categories in a way that implied potential items for model testing.  

**Figure 2**
The purpose of phase 2 was to test the meta-model from phase 1 using a mixed-methods experimental approach (MM-EXP). To compare high-quality and low-quality online tutoring sessions, we measured both quantitative and qualitative data from additional archived session transcripts and further developed a theoretical model that addressed the overarching question. The research questions were:

Research Question 1: What theoretical model, grounded in qualitative and quantitative data, reveals the hallmarks of one-on-one high quality online tutoring?

Research Question 2: What are the latent variables and factors of each in a model of online tutoring quality?

Research Question 3: Does the model support significant differences between variables from rated low-quality and rated high-quality sessions?

Item Development

Based on DeVellis and Thorpe’s scale development process (2022), items were generated based on the models built during phase one. We used anecdotal questions and statements from previously viewed transcripts to generate an item pool aligned to factors, sub-categories, and
categories. During item writing adjudication, we collapsed several categorical ideas into single items. We produced 21 items in all, one item per sub-category for conservation of instrument size. We chose items that could be answered with a few, simple responses, such as not applicable, no, inferable, and yes. A sample of this alignment can be seen in Table 1.

### Table 1

**Sample of construct and item development alignment**

<table>
<thead>
<tr>
<th>Category</th>
<th>Sub-category</th>
<th>Factor</th>
<th>Anecdotal evidence</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communication</td>
<td>Interactive</td>
<td>Engaging</td>
<td>“How can I help?” “What should we tackle first?” “Do you want to work on the board?”</td>
<td>Is there evidence that the tutor created an experience that engaged the learner?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Promptly responds</td>
<td>No notable delays between tutor to student responses</td>
<td>Is there evidence of good pacing by the tutor without unnecessary delays between responses?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Supportive</td>
<td>“I understand. Unfamiliar formats can feel stressful at times.” “You got it! Nice job!”</td>
<td>Is there evidence that the tutor was supportive of the learner in an appropriately positive manner?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Encouraging</td>
<td>“This looks good so far. Let’s keep going.” “Do you want to try another?”</td>
<td>Is there evidence that the tutor was encouraging the learner in an appropriate manner?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Clarifying (closure)</td>
<td>“Try the next problem and log back on if you need to.”</td>
<td>Is there evidence that the tutor ended the session appropriately?</td>
</tr>
</tbody>
</table>

Two additional individuals, subject matter experts familiar with online tutoring and session transcript reading, were asked to assess transcripts using the developed scale. Both shared experiences using the instrument, the process, and ease of use.

A Principal Component Analysis (PCA) was conducted using IBM’s SPSS software version 24. When we discovered items that were considered too highly correlated, we removed two items (Item 11- Know-Content-Error Free) and (Item 9-Know-Tech-Policies) due to redundancy and hindering of maximum independence of factors. A Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was 0.846, greater than 0.5 indicating sample size was sufficient. The Bartlett’s Test of Sphericity was statistically significant ($p = 0.000$) indicating that there were at least one pair of correlated factors. Communalities ranged from 0.546 to 0.922. The total variance matrix, echoed by the scree plot, indicated there were 5 factors with eigenvalues above 1 out of the remaining 19 items, explaining 74% of the model. Three factors had eigenvalues between 1 and 2, and two factors were greater than 2. The component correlation matrix revealed factor relationships that were orthogonal.

Exploratory Factor Analysis (EFA) was performed to further refine the model. Using Varimax settings, the rotated component matrix was calculated. This led to removing a few more items (Item 3-Comm-Style-Approachable and Item 4-Comm-Style-Refrain from personal info) based on low communalities. Further items were removed in components containing only one item (Item 17 – Instruct – Plans – Approach) Choosing three fixed factors The Rotated Component
Matrix was used to identify which items belonged to which component. All factor loadings were above 0.500 but a few did load on more than one component.

Next up was a Cronbach’s alpha test of reliability for factors/items and components. Using SPSS again, we found Cronbach alpha’s scores of 0.897, 0.824, 0.916, respectively. Based on this we removed 3 items (Item 12-Know-Content-Pacing, Item 14-Instruct-Assess-Goal stated, and Item 7-Comm-Interactive-Departed warmly). The overall Cronbach reliability yielded for the remaining 12 factors was 0.927, an excellent reliability score.

**Table 2**

*Items classified based on Rotated Component Matrix using Varimax Rotation*

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Item 5</td>
<td>Item 2</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Item 6</td>
<td>Item 8</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Item 20</td>
<td>Item 13</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Item 21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Using IBM’s Amos software, a Confirmatory Factor Analysis was conducted. This involved building a structural equation model of observed variables and their latent variables. It also involved finding if the model was a good fit and using modification indices to help determine where it may not be. In this process, 4 more items were removed (Item 10 – Know – Content – Refuse/Transfer, Item 17 - Instruct - Plans – Approach Change, Item 20 - Instruct - Clar&Adj – Learner Efficacy, and Item 21 - Instruct - Clar&Adj - Sound Ending) and the final model contained 8 factors, seen in Figure 3.

The final measurement model consisted of three components, accounting for 82% of the data. Assessment of KMO (0.822), Bartlett’s ($p = 0$), and Communalties (0.727-0.929) were all good. The model was evaluated using several measures of test fitness using the Amos software. The comparative fit index (CFI) and the Tucker Lewis Index (TLI) yielded 0.978 and 0.964, respectively, both considered excellent fit. The root mean square of error approximation (RMSEA) was 0.77, an acceptable fit in this case (Xia & Yang, 2019).

**Figure 3**

*Structural equation model of the hallmarks of high-quality online tutoring*
Research Question 2: What are the latent variables and factors of each in a model of high quality online tutoring?

The Hallmarks of High-Quality Tutoring

Evidence of Student-Centered Communication:

**Engaging** – Tutor demonstrates attempts to engage the learner by creating an interactive environment of give-and-take, prompts and responses, with a natural flow of responses by both parties; no refusal to help, no transfer of session, no unnecessary delays nor disconnections.

**Supportive** – Tutors demonstrate supportive, often positive responses, that pertain to the development and realization of learning or the learning process by expressing explicit acknowledgement (and sometimes empathy) to learner responses regularly throughout the session; no elaborate discussions of unrelated nor personal information, no pushing forward without acknowledging learner response.

Evidence of Instructional Strategies:

**Accesses prior knowledge** – Tutor asks the learner about related, already-established information in regard to the problem or assignment provided that is appropriate to learner level and topic; no excessive nor extraneous background probing. Tutor accesses only information related to moving forward in addressing the request on hand. (This may look different if the learner is seeking a general review of topics for an upcoming exam, for example).

**Asks clarifying questions** – Within and often throughout the session, the tutor asks questions to reveal information from the learner as opposed to making assumptions about what the learner should understand from the process. These are more often to occur earlier in the session in developing a consensual understanding of the needs, expectations, and outcomes for the session.
**Scaffold learning** – Tutor mostly uses inquiry to guide learning or provides information in short statements, whether generated by inquiry through questioning or chunking of answer-giving as opposed to lengthier narrative information giving. Examples include providing a diagram to build or enrich understanding of a topic, a reminder where APA rules can be easily located, or reminding a learner of a formula if it cannot be generated through inquiry or from using class notes.

**Ensures understanding** – Tutor regularly checks-in to ensure that a learner is remaining actively engaged throughout the session by asking check-in questions. Examples include, “Does this make sense?”, “Do you understand what you just wrote on the board?”, “Do you want to try an example?”.

Evidence of Tutor Efficacy:

As subject matter experts, tutors are expected to know procedural and conceptual knowledge of the subjects they tutor. This is an expectation by learners when requesting tutors. Therefore, occurrences and invariance were high in both low-quality and high-quality sessions. Therefore, as a factor, this could not be used to differentiate between sessions. However, knowledge of the platform and tools and more importantly, a tutor’s familiarity and flow utilizing those tools, had an impact on learner satisfaction and subsequent rating.

**Use of Modalities** – Tutor engages more than text in sessions, such as the board, shared image, graph paper, coding editor, or some other visual modality. This can also include file sharing and use of audio, where applicable.

**Familiar with Platform Tools** – Tutor demonstrates familiarity with platform, procedures, tools, and their usage as demonstrated in session; no delays or misunderstandings in using tools by tutors, explanations given to learners who may be struggling to use a tool and is resolved in high-quality tutoring sessions.

**Phase 3: Model testing**

The purpose of phase 3 was to test the meta-model from phase 2 using experimental methods. The research question addressed was: Are there differences between variables from the Hallmarks of High-Quality Tutoring model between low-quality and high-quality sessions?

**The Sample**

The Princeton Review’s Tutor.com provided archived transcripts of online tutoring sessions that occurred between 2017-2019. We started with 150 archived transcripts from math, science, and English sessions conducted. Of those, 119 met the requirements (no audio sessions, all over ten minutes, all conducted with a single tutor). Of these, 62% were high-quality sessions based on high ratings (5), positive student comments, and positive tutor comments. The remaining 38% were low quality, based on low ratings (1), negative student comments and neutral or negative tutor comments. Transcripts represented 31 Algebra, 5 Pre-Calculus, 30 Biology, 4
Chemistry, 39 English, and 10 ELL (English Language Learning) sessions. We were interested in a model that was generalizable across subjects and disciplines.

An independent-samples Student t-test was conducted to compare high and low quality of tutoring sessions by student centered, instructional, and tutor efficacy variables. The results of Levene’s test were significant and the null hypothesis rejected \((\mu \neq \mu_0)\) that variances were not equal. In the overall sample, the skew indexes for factors ranged from 0.648 to 0.222 and kurtosis indexes ranged from −0.724 to 0.440. The means and standard deviations for the factors for the full sample, as well as the latent variables: Student Centered, Instructional, and Efficacy, and Total are reported in Table 3.

**Table 3.**

**Means and standard deviations of the latent variables**

<table>
<thead>
<tr>
<th>Latent variable</th>
<th>Overall sample</th>
<th>High quality sample</th>
<th>Low quality sample</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Student Centered</strong></td>
<td>(M = 1.42) (SD = .835)</td>
<td>(M = 1.86) (SD = .375)</td>
<td>(M = .711) (SD = .895)</td>
</tr>
<tr>
<td><strong>Instructional</strong></td>
<td>(M = 1.09) (SD = .751)</td>
<td>(M = 1.34) (SD = .548)</td>
<td>(M = .694) (SD = .868)</td>
</tr>
<tr>
<td><strong>Efficacy</strong></td>
<td>(M = .861) (SD = .325)</td>
<td>(M = .966) (SD = .126)</td>
<td>(M = .689) (SD = .456)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>(M = .115) (SD = .294)</td>
<td>(M = .692) (SD = .857)</td>
<td></td>
</tr>
</tbody>
</table>

There was a significant difference in the scores for **Student Centered** between high-quality \((M = 1.86\) \(SD = .375\)) and low-quality sessions \((M = .711\) \(SD = .895\)), \(t(117) = -0.9729, p = .000\); **Instructional** between high-quality \((M = 1.34\) \(SD = .548\)) and low-quality sessions \((M = .694\) \(SD = .868\)), \(t(117) = -4.962, p = .000\); and, in **Tutor Efficacy** between high-quality \((M = .966\) \(SD = .126\)) and low-quality sessions \((M = .689\) \(SD = .456\)), \(t(117) = -4.942, p = .000\). Additionally, an overall comparison of the means between the two groups yielded a significant difference, \(t(117) = -5.766, p = .000\). The results suggest that when tutors are student-centered, instructional, and efficacious, they impact the quality of online tutoring sessions.

**Discussion**

**Using the TPACK lens**

The primary question of this study was, *what theoretical model, grounded in qualitative and quantitative data, reveals the hallmarks of one-on-one high quality online tutoring?* During the framing of this study and review of the literature, a lens using the TPACK theory led to several questions for consideration. Many of the research-based considerations led to item writing. Many of those items were removed during factor analysis. Analyzing all of the original items provides responses to earlier questions. Using the information collected during phase 3 \((n = 119)\), the findings for each are discussed below.
**Technological knowledge**

*How valuable is a tutor’s familiarity with platform tools?* Is knowing how to transition between capabilities and tools, incorporating multiple modalities into the tutoring session, and overcoming difficulty the learner may be having with tools and/or setting, valuable to learners in *high quality online tutoring*? From this data set, there were clear indications that all of these considerations were valuable to learners. There were several incidents found in low-quality sessions where a tutor struggled with using or explaining file sharing, for example, in which the tutor struggled to assist. This extended the length of the session which likely contributed to lower student satisfaction as evidenced by ratings and comments. Similarly, a connection was found between a tutor’s familiarity with policies and procedures, such as engaging topics unrelated to the problem or topic posed. This was only observed in low quality sessions. In higher quality sessions, use of chat accompanied by another modality, use or writing on the board, for example, was evident in almost every high-quality session. See figure 4.

**Figure 4**

*Comparison of quality in technological knowledge of tutor*

![Tutor Technological Knowledge](image)

**Content knowledge**

*How valuable is the tutor’s content and procedural knowledge?* Is accepting student requests when the question/problem is revealed, using accurate vocabulary, rules, formulae, and references/resources, and assisting over transferring to another tutor, valuable to learners in *high quality online tutoring*? There was evidence that many of these considerations were valuable to learners. A demonstration of a lack of content knowledge was interpreted when a tutor stated that they were not familiar with a topic in the given subject area or claimed a lack of expertise or transferred the session; did not engage nor embrace the learner and problem presented. This was only observed in sessions that were deemed lower quality.
One realization that emerged is that learners are generally unable to verify content and procedural knowledge until after the session when one applies their newly gained knowledge in class, for example. For this reason, this item did not vary much between high and low-quality tutoring interactions. All sessions examined were first filtered for length over 10 minutes so in many cases, when this late transferring occurred, it was deemed wasteful and frustrating to the learner to have been greeted, shared the information (problem or paper) and then told they could not be helped. Another construct that emerged was that of tutor efficacy. None of the higher-quality sessions in this review demonstrated instances of apparent errors, apologies or utterances of doubt. All tutors in these sessions sounded confident and capable in their subject area. There were, however, incidences in lower-quality session transcripts. See figure 5.

Figure 5

Comparison of quality in content knowledge of tutor

![Comparison of quality in content knowledge of tutor](chart.png)

**Pedagogical knowledge**

**How valuable is the tutor’s pedagogical knowledge?** Is organization, engagement, interaction, availability, communication, professionalism, and conscientiousness valuable to learners in high quality online tutoring? From a review of the data, there was an abundance of evidence to indicate that these considerations were valuable to learners and present in high-quality sessions. Interactions between tutor and student seemed organized, professional, purposeful and mindful of each other’s responses and inquiries. Demonstrations of a back-and-forth flow of utterances in high quality sessions and a sometimes lack of understanding and flow in the exchanges between tutor and student in low quality sessions were gleaned from chat logs. Goal setting or iterating expectations were more common in high quality sessions as was demonstrating a purposeful approach, changing that approach when flow was disjointed to
disrupted, and encouraging the learner to continue examining and building upon their mental schema. Interestingly, encouragement and support were mainly positive but not always effective. In some lower-quality sessions, the tutor seemed to overuse emojis and use inflated positive phrases when minor steps were accomplished. This may be an area for further research given this data set only incorporated students enrolled in higher education. In some sessions where the learner failed to answer the tutor’s question correctly, the tutor did not respond positively but was still deemed as encouraging. In others, a tutor’s succinct nature was observed by raters as being interpreted as tone by students. For this reason, we changed the term *positive* to *supportive* regardless of if the statement was positive or negative by itself. See figure 6.

**Figure 6**

*Comparison of quality in tutor approach*

![](image)

**How valuable is inquiry?** Is the questioning of students to engage, explore, explain, elaborate, and evaluate valuable to learners in high quality online tutoring? Is inquiry-based questioning over answer-giving valuable to learners in high quality online tutoring? Indications from coding of quotes between learner and tutor showed a variety of inquiry-based questions and comments that resembled attempts to engage, explore, explain, elaborate and/or evaluate. Not all sessions considered high-quality revealed every step of the 5E model, but observations were recorded when observed. From comparison, there was evidence of giving too much of an answer to the learner with far less engagement in the sessions deemed low quality. There were occurrences noted of answer-giving in higher quality sessions as well with notable differences; short bursts or chunks of answer-giving seemed to function as enrichment to conceptual understanding. See figure 7.

**Figure 7**

*Comparison of quality in use of inquiry by tutor*
How valuable is a tutor’s ability to create an engaging and satisfying interaction? Is encouragement and positive utterances valuable to learners in high quality online tutoring? Our findings resonated with studies that inferred quality is influenced by positive utterances (Chen et al., 2019). We noted the more utterances, the more likely inquiry questions were noted but would suggest further study in this area. One potential measure would be the number of back-and-forth utterances. While positive utterances were noted more often in high-quality sessions, support in these sessions was also observed as indifferent or even negative but still successful. Those utterances considered indifferent or negative were still supportive. Being supportive of learning and encouragement seemed to be key to continuing the session and learning. In another area of inquiry, questions asked by tutors included checking in with what seemed like attempts or served to clarify or ensure student understanding. Support for this stems from Frey, Fisher, and Alamode’s (2021) work, How Tutoring Works, which stated that tutoring works when teachers use clarity to build relevance. See figure 8.

Figure 8

Comparison of quality by tutor interactions
How valuable is timeliness and utterances? Are utterances, length of session, pacing in session, and length between responses valuable to learners in high quality online tutoring? Related to utterances was how much time passed between them, more delays in responding occurred in low quality sessions. High-quality sessions were more likely to demonstrate appropriate pacing between utterances. This included appropriately ending tutoring sessions in a naturally or educationally sound manner. Low-quality sessions were more likely to contain longer and shorter time between utterances from student and tutor. Low-quality sessions were also more likely to abruptly end with either the student or tutor disconnecting without explanation or the tutor stating they had to end the session (shift was ending, unable to help, lack of expertise). In both types of sessions, warmly greeting students and warmly departing was evident. This is one reason those items did not hold up in factor analyses. See figure 9.

Figure 9

Comparison of quality by tutor utterances
Conclusion

Ample studies continue to support the benefits of online tutoring. This study, however, aimed to address the evaluation of effective tutoring, with implications for those tasked with preventing ineffective tutoring by understanding what makes a high-quality versus a low-quality tutoring session. Research Question 1 was: What theoretical model, grounded in qualitative and quantitative data, reveals the hallmarks of one-on-one high-quality online tutoring?

From a reiterative MM-GT design process, the final constructs and items are categorized into three latent variables: tutor efficacy, student centered, and instructional. Therefore, the hallmarks of a high-quality online tutoring session are: (a) Tutor efficacy of platform technology and tools; (b) Use of multiple modalities when applicable; (c) Engagement evidenced as multiple and appropriately paced interactions; (d) Inquiry and incorporation of student’s prior knowledge; (e) Support as encouragement of student efforts; and (f) Ensuring understanding through clarifying and checking-in with student.

Figure 10

Hallmarks of High-Quality Tutoring
To explore if these hallmarks could distinguish between low and high-quality online tutoring, approximately 150 session transcripts were analyzed quantitatively. The overall comparison of means between the two groups was performed. While the majority of the 150 session transcripts under review yielded high quality sessions, comparing the means yielded a significant difference, $t(117) = -5.766$, $p = .000$. Differences were seen at each factor level (hallmark) and latent variable. The results suggest that when tutors are student-centered, instructional, and efficacious, they positively impact the quality of online tutoring sessions. Following additional coding of qualitative information, enriched definitions were made for each of the hallmarks.

Evidence of Tutor Efficacy of online tutoring platform and tools:

**Manifests Platform Capabilities** – In high quality sessions, tutors demonstrate familiarity with platform, procedures, tools, and usage as demonstrated in session. There are no misunderstandings nor delays in using tools by tutors and explanations are provided to learners who may be struggling to use a tool. Any difficulty is resolved as a result.

**Use of Multiple Modalities** – In high quality sessions, tutors engage with learners in multiple ways using multiple tools. This includes use of a board, sharing images or diagrams, working on graph paper, using a coding editor, or some other visual modality including and marking up of images or files with annotations. These are in addition to an area for texting. This can also include audio or video, which was eliminated from this data set to avoid potentially confounding variables.

Evidence of Instruction:

**Accesses and builds upon a student’s prior knowledge** – Tutors begin by reviewing a posted statement or problem and greeting the student. In high quality sessions, tutors are more likely to ask the learner about established or associated information, revealing mental schema a student had concerning the objective. This is observed in contrast to a tutor that would review the problem and then lead the learner through lecturing about the solving of the problem. While accessing prior knowledge, the tutor should refrain from excessive nor extraneous background probing, acclimate to learner level and then build upon the provided information in attempts to move the interactions forward.

**Inquires using open-ended questions** – In high quality sessions, effective tutors use inquiry to guide learning for the student. Inquiry is used to scaffold the constructivist building of generating connections and “realizations”. In some cases, tutors may provide information, as short statements as opposed to lengthier narratives. Answer-giving and lecturing by tutors are more commonly found in low quality sessions. When short bursts of information are found as answer-giving in high quality sessions, these are perceived more likely to be used to enrich to the topic being discussed. Many inquiry-based questions can be categorized or associated with the 5E model of inquiry-based learning such as: asking to elaborate or explain, asking to engage or explore, and asking to evaluate.

**Clarifies to ensure understanding** – Highly effective tutors ask questions to reveal their assumptions about what they think the student knows or was learning through the process. This seemed to be a critical strategy for the successful tutor. By clarifying
understanding, the tutor continues with current instructional approach or changes their approach as needed. Without this information, the tutor might be oblivious to learner frustration, confusion, and/or disappointment. This can be seen in low-quality sessions when students repeat the term, “okay” without being able to move forward in solving the problem. Some students disconnected or will voice concern that they are not being helped or that they are getting [more] confused.

**Connecting and maintaining engagement** – In high-quality sessions, tutors will connect to the learner. The repour between tutor and student will seem to flow better. Tutors that check-in with learners maintain a momentum of active engagement. Tutors’ questions should relate to the process in this regard. Examples of these types of questions include, “Does this make sense?”, “Do you understand what we did on the board?” “Do you want to try an example?”, “Shall we continue?”.  

Evidence of Student-Centered communication:

**Engagement through multiple and appropriately paced interactions** – In high-quality sessions, tutors more frequently demonstrate an attempt to engage students and embrace learning. These tutors seem to create an interactive environment of give-and-take, prompts and responses, with a natural flow of timely responses. In low-quality sessions, tutors are more likely unable to help or understand the learner or the problem. Tutors in these sessions claim the topic is out of topic (when such topics are appropriate for the subject area), transfer the session to another tutor (sometimes repeatedly), decline to use a modality the student requested, had long delays between responses, or disconnected suddenly.

**Support to encourage learning by tutor** – Tutors in high quality sessions respond to learner efforts and utterances. Often, tutors have positive responses, or smiling emojis, and are empathetic, where appropriate. Tutors use supportive encouragement during the developing learning process or when learners display realization of concepts. In high quality sessions, this does not include elaborate discussions of unrelated nor personal information. In low quality sessions, support was sometimes evident as overinflated or completely absent, or lacking an acknowledgement of the learner’s response.

**Conclusion**

These hallmarks were found rooted in traditional constructs of learning. Strategies that enhance learning, self-efficacy, and student success all incorporate active learning and online tutoring is inherently active (Bransford, Brown, & Cocking, 2000; Jansson et al., 2021). The Hallmarks resonate with constructive cognition (Piaget’s constructivism) for the rearrangement, assimilation, and changes in an individual’s mental schema and support social interactions and language used between learner and tutor as a shared learning experience (Vygotsky’s social constructivism). These Hallmarks consist of: tutor efficacy of platform technology and tools, use of multiple modalities when applicable, engagement evidenced by multiple and appropriately paced interactions, incorporation of student’s prior knowledge, tutor use of inquiry through use of open-ended questions, student support as encouragement of efforts by tutor, clarifying information by tutor to ensure student understanding, and checking-in and connecting with student to ensure engagement and should be further studied. Implications for the use of
hallmarks include hiring, training, and/or evaluating tutors by those who oversee tutoring programs. By ensuring tutor quality for each and every session, online learning experiences can reach the potential for every learner to further understanding and reach success. Our objective was to discover the hallmarks of high quality online tutoring sessions so those tasked with overseeing tutoring programs can offer effective session for all students, regardless of subject. Tutoring is not lecturing nor answer-giving. To foster academic independence by students, a tutor should conduct sessions that meet the hallmarks of a high-quality session.

**Recommendations for further research**

Some areas for potential research include: (a) further exploration of the inquiry-based 5E model of instructional strategies as found in tutoring interactions, (b) exploring other grade levels, such as K-12 learners and their online tutoring experiences, (c) exploring differences between subjects, essay writing as compared to algebra, for example, and (d) relationships between high quality and low-quality tutoring and program success and/or retention.
References


