Instructional Science
https://doi.org/10.1007/s11251-022-09604-6

ORIGINAL RESEARCH

Examining relations between math anxiety, prior knowledge, hint usage, and performance of math equivalence in two different online learning contexts

Alisionna Iannacchione[1] · Erin Ottmar[1] · Vy Ngo[1] · Craig A. Mason[2] · Jenny Yun-Chen Chan[1,3] · Hannah Smith[1] · Kathryn C. Drzewiecki[1] · Stacy T. Shaw[1]

Received: 31 January 2021 / Accepted: 13 October 2022

Abstract
Prior research has shown negative relations between math anxiety and math performance. We posit that one potential pathway through which math anxiety influences performance of math equivalencies is through help seeking behavior during learning. Here, we examine whether middle school students' behavior, specifically the frequency of hint requests, within educational technologies mediates the association between math anxiety and performance of math equivalence. Students completed a pretest measuring their performance of math equivalence and math anxiety prior to the intervention, and a posttest measuring their performance of math equivalence. We examine mediation in two online math learning technologies: From Here to There (FH2T) and ASSISTments. In both FH2T and ASSISTments, students can request hints that provide just-in-time support during problem solving. We examined whether the frequency of hint requests mediates the effects of math anxiety on performance in both conditions. Using multi-group mediation analyses, we found that math anxiety was not a predictor of hint usage in either condition when controlling for pretest performance. Further, we found that students with lower performance at the pretest used more hints in the problem set condition, and using more hints was associated with lower performance of math equivalence at the posttest. This relation was not significant in the FH2T condition, suggesting a fundamental difference in hint usage between the two technologies. These findings have implications for designing educational technologies that simultaneously promote math performance and productive help seeking behaviors in middle school students.

Keywords Math anxiety · Technology · Hints · Math equivalence

✉ Alisionna Iannacchione
amiannacchione@wpi.edu

1 Social Sciences and Policy Studies, Worcester Polytechnic Institute, Unity Hall 320, 100 Institute Road, Worcester, MA 01609, USA

2 School of Learning and Teaching, University of Maine, Orono, ME, USA

3 Department of Early Childhood Education, The Education University of Hong Kong, Tai Po, Hong Kong SAR, China

Actively seeking help can be especially challenging for students with high math anxiety, who may experience intense fear and worry at the thought of doing math (Lyons & Beilock, 2012) or seeing a math formula (Pizzie & Kraemer, 2017). Also, students often struggle to engage with mathematics and ask for help when they need it (Newman, 2008). This could be due to lack of self-efficacy, classroom structure, social factors, and other reasons (Ryan et al., 1998). High math anxiety is linked with poorer performance and math achievement (Foley et al., 2017), and students who are math anxious often show patterns of math avoidance and seek less help (Choe et al., 2019; Ramirez et al., 2018), which in turn prevents them from improving their math performance.

Emerging educational technologies, where students work their way through online lessons or math problem sets that include readily available hints may help mitigate some of these help-seeking and avoidance issues by design (Razzaq & Heffernan, 2010). Many educational technologies have integrated features that encourage students to seek help. For example, students in an online environment can simply click a button to request and receive a hint or help on a problem without any form of social pressure (e.g. Anderson et al., 1995; Heffernan & Heffernan, 2014), whereas students in traditional classrooms may experience anxiety when raising their hand to ask their teacher a question or expressing confusion in front of their peers (e.g. Karabenick & Knapp, 1991; Newman, 1990). Such affordances in educational technologies can transform an act of help-seeking that traditionally is social and potentially risky (interrupting a lesson to ask a question) with something that has a lower perceived risk and is less visible to others in the classroom (clicking a button). In addition to providing a safe space for students to request help during problem-solving, many educational technologies record student actions and provide educators and researchers the information to explore and potentially improve students' help seeking behavior (e.g. Kehrer et al., 2013; Razzaq & Heffernan, 2010).

This unique affordance of hints and granular data logged within educational technologies provides the opportunity for researchers to explore how math anxiety, help-seeking behavior, and math achievement relate to each other. While much research on educational technology compares learning in one system to traditional business-as-usual classrooms (Kelly et al., 2013; Singh et al., 2011), direct comparisons between different educational technologies can provide valuable insights. Specifically, these comparisons can reveal how differences in technology design may impact student learning. In this paper, we explore the relations between prior knowledge, math anxiety, and hint usage, and how these relations impact performance of math equivalence when students receive an intervention with one of two different educational technology tools: From Here to There! (FH2T; (Ottmar et al., 2015a, b) or problem sets in ASSISTments (Heffernan & Heffernan, 2014). These two technologies both have demonstrated efficacy for improving math learning; however, they differ significantly in design, including the function, presentation, and impacts of hints available on these two tools. FH2T presents hints in a playful, game-based environment that encourages flexible problem solving in which there is no score, and the hints give the student a first step to an answer path; whereas the problem set system in ASSISTments offers a partial credit system in which points are deducted in percentage as students request hints and these hints follow one path to the correct answer. Because of these distinctions, we expect to see differences in how students' use of hints impacts their posttest performance between the two educational technologies. Further, we anticipate that the design and function of hints play a role in mediating the relations between student's anxiety and performance of math equivalence. Specifically, we predict that the game-based hints in FH2T may encourage effective help seeking behaviors and be less anxiety inducing to mitigate the negative effect of math anxiety on performance.

Math anxiety

Math anxiety refers to fear or apprehension related to engaging with math (Ashcraft, 2002). For students who are math anxious, common educational activities such as completing homework problems, taking math assessments, demonstrating problems on the board, asking the teacher for help, or even just seeing a math problem can trigger intense worries and heightened physiological responses (Lyons & Beilock, 2012; Pizzie & Kraemer, 2017). Across numerous studies, high math anxiety has been found to be associated with poorer math performance (Foley et al., 2017; Ma, 1999; Wu et al., 2012). Some research has found evidence for a reduced competency account, which contends that students develop math anxiety because they lack the foundational math skills necessary to perform well (e.g., Maloney et al., 2010). This development of math anxiety may be driven by avoidance of math and poor help-seeking behavior (Ramirez et al., 2018).

There has also been research to support that there may be differing effects of math anxiety on performance based on gender. One study found that the negative correlation between math anxiety and arithmetic performance was shown exclusively in girls (Van Mier, et al., 2019) while another study showed that even though girls typically have higher anxiety, boys suffer more negative effects on math performance due to math anxiety (Szczygiel, 2020).

While the effects of math anxiety in the classroom are well-known, it remains unclear how these effects translate into the online learning sphere and how help-seeking in those online learning tools plays a role. One study comparing a game-based learning technology, a more traditional online learning technology through eBooks, and a business-as-usual condition with classroom instruction found that students in the game-based technology showed the most improvement in math performance. However, no significant difference in math anxiety from pretest to posttest was found between the three conditions (Hung et al., 2014).

As online math learning platforms, both game-based and answer-based, become more common, exploring the potential relations between math anxiety and learning in different platforms is critical for improving future design of online systems. Further, many teachers are increasingly using online learning platforms (e.g. FH2T, ASSISTments) to both facilitate engagement as well as to enhance student learning. Despite the increasing popularity of using technology in the classroom, there is a scarcity of research on using technology to relieve the effects of math anxiety (Iossi, 2007). Some research suggests that students may experience lower levels of anxiety when given the opportunity to complete math problems online; however, this relation may depend on the level of anxiety a student typically has in class and during exams (Stowell & Bennett, 2010). Further, it is possible that the inherent designs of educational technologies—such as how students seek help in these systems—can differentially alter the impact of math anxiety on later performance (Beilock & Maloney, 2015).

Help-seeking behavior in classrooms

Academic help-seeking behavior can be both formal and informal (Karabenick & Berger, 2013; Karabenick & Knapp, 1988). In classroom settings, academic help seeking takes many forms. Students who are struggling may seek help in the form of asking questions during class or contacting the instructor outside of class, or by turning to classmates, peers, or friends for questions and support. In many technology-based learning environments, students can request and receive immediate help during learning and problem solving. Specifically, with the rapid advancement of educational technologies, online learning platforms offer several different types of on-demand support to students, ranging from context-specific hints (Anderson et al., 1995; Stamper et al., 2011) and worked examples, to general background support materials such as online glossaries and videos (Aleven & Koedinger, 2000; Whitehill & Seltzer, 2016).

Overall, effective help-seeking behavior, both formal and informal, has been shown to improve learning and increase performance in school through hints that prompt students to employ a specific strategy or showing an example of a strategy being used correctly (Renkl, 2002; Wood & Wood, 1999). However, there is growing evidence that students may *not* always exhibit effective help-seeking behaviors that lead to gains in learning and performance, and that students seek and react to help differently depending on their prior knowledge, achievement goals, and self-perceptions for a variety of reasons (Aleven et al., 2003). For instance, students are less likely to seek help if they already feel vulnerable about their knowledge or ability in a subject (Karabenick & Knapp, 1991). Further, students perceiving help-seeking as a threat to their self-esteem may be less likely to seek help formally in classrooms (Karabenick & Knapp, 1991). Due to the perceived psychological risks of help seeking by students, a recent study found that students were more likely to seek help in private instead of publicly in the classroom (Peeters et al., 2020). Thus, students' self-perceptions or psychological risk, such as concern about self-esteem, can greatly impact their likelihood to engage in help-seeking behaviors in classrooms.

Help-seeking behavior in online educational technologies

The relations between students' affective factors and their help-seeking behaviors might differ in online educational environments, in that students may not ask for hints due to social pressure in the classroom but may ask for hints from online systems. While online learning technologies allow students to receive critical help and resources without the social risks of help seeking, there are drawbacks to this. Previous studies have shown that when using educational technologies, some students may exhibit help-seeking strategies that are not conducive to learning. For instance, students often ignore the help facilities, such as hints, on the tutoring system or use them in ways that are not likely to promote learning (e.g., game the system by requesting all hints including the correct answer), suggesting that ineffective hint use may diminish the benefit of hints in the tutoring system (Aleven et al., 2003). Further, Aleven and Koedinger (2001) found that students with lower prior knowledge requested hints more frequently compared to students with high prior knowledge, yet requesting more hints was associated with lower math performance, suggesting that any potential benefit from hints was not sufficient for improving math performance. In educational technologies, when the student is responsible for discerning their need for help, these ineffective help-seeking behaviors can play a great role in limiting their learning (Aleven & Koedinger, 2000, 2001; Aleven et al., 2003).

Plenty of work has since explored how help can be designed and incorporated in educational technologies to mitigate the negative effects of affective factors and encourage students' productive help-seeking behaviors. For instance, ASSISTments allows teachers and researchers to flexibly implement hints and scaffolding in the form of immediate or delayed support. A study conducted on ASSISTments (Kehrer et al., 2013) found that students who completed math homework with correctness feedback immediately available learned more compared to students who only received correctness feedback at the end of the assignment. Another study that used ASSISTments (Razzaq & Heffernan, 2010) also found that when on-demand hints are available, middle school students who asked for more hints performed better on algebra assignments. These results suggest that when hints are available on demand in educational technologies, students may be more likely to engage in help-seeking behaviors that lead to learning gains. However, the benefits of on-demand hints are still uncertain; there is varying evidence of when hints are helpful and to what extent the context in which hints are presented influences learning (Price et al., 2017). For example, an initial study first showed that students improved their math performance when on-demand hints were available, yet a replication study later did not find performance gains (Inventado et al., 2018).

Recent studies have also shown that design mechanisms in the help-seeking features of math computer assisted learning systems may help mitigate poor help seeking behaviors. Students demonstrate more learning gains when their hint use was regulated by the system and rejected for a period of time if it was deemed to be overused as a way to encourage students to attempt on their own compared to those that were able to ask for hints at all times without regulation (Chou & Chang, 2021). Another pilot study found evidence for learning gains in a data driven hint display that responds differently based on a student's help seeking tendency (i.e., help avoidance or help-abuse), by either encouraging or reminding students of the presence of hints or not allowing hints to become available for a period of time (Marwan et al., 2020).

As evident in prior studies, the effects of student factors, such as self-perceptions and math anxiety, on their help-seeking behaviors are unclear and complex, especially in online learning environments. Such complicated relations have presented a particular challenge for researchers and educators to determine how help should be designed in educational technologies. Given that productive help-seeking can positively impact students' performance and learning, it is critical to understand these relations and to develop systems that provide students with appropriate hints and encourage effective help seeking behaviors.

The present study

In the present study, we examine how the impact of math anxiety on performance in mathematics equivalence differs between two educational technologies and how the use of hints within these educational technology tools may influence the association between math anxiety and math performance. Further, we explore the influence of students' prior knowledge on the relations between math anxiety, hint requests, and performance of math equivalence. We utilize data collected from a classroom study testing the impacts of two educational technologies—FH2T and problems sets in ASSISTments—on equivalence understanding.

FH2T (https://graspablemath.com/projects/fh2t) is an online, self-paced math game in which students dynamically manipulate equations and expressions to match specified goals (Ottmar et al., 2015a, b). ASSISTments is a problem-based educational technology where students solve traditional textbook problems and receive immediate correctness feedback on their answers (Heffernan & Heffernan, 2014). Both systems have been shown to be effective at improving students' math performance (Chan, et al., 2021; Hulse et al., 2019; Roschelle et al., 2016). In both systems, students can request hints during problem-solving, and they need to solve the problem to proceed in their learning activity. However, these technologies differ in many aspects, including how the hints are designed. In FH2T, one hint is available per problem that provides students with a cue about a possible pathway to solve the math puzzle. The hint encourages students to think about problem solving processes but does not penalize students for using it. In the problem set condition, three hints are available per problem: a conceptual hint, a worked example of a similar problem, and a worked example of the given problem. Students can request the correct answer to the problem after the three hints. The students can request hints at any time, but the system deducts points for each hint requested. In comparing these two technologies, we are essentially examining the effects of the ways in which hints and immediate feedback are presented on math anxiety and equivalence understanding, an aspect of algebra learning.

Because of the various implications of math anxiety, including its potential role of math anxiety in help-seeking behavior, we investigate how students demonstrating help-seeking behaviors in the form of using hints in online technologies may mediate the effects of math anxiety on performance of equivalence. We have conducted a randomized controlled trial with a pretest measuring math anxiety and performance, four sessions of technology intervention in FH2T or a problem set condition, and a posttest measuring math performance of equivalence. In the current study, we examine the relations between pretest math performance of equivalence, math anxiety, hint usage, and posttest math performance in the context of FH2T and the problem set condition in ASSISTments. Further, we explore how these two technologies may differentially impact the association between math anxiety and performance through students' hint requests within the technologies. Figure 1 shows the conceptual model that we are testing in this study.

Our hypotheses are as follow:

1. Higher pretest math anxiety would be associated with lower posttest equivalence performance (the direct path C; Foley et al., 2017; Ma, 1999; Wu et al., 2012).
2. Students with higher math anxiety would request more hints in the online environment as the online vs. classroom context might involve fewer negative social consequences (path A; Karabenick & Knapp, 1991; Newman, 1990), and more hint requests would be related to lower posttest equivalence performance (path B; Aleven & Koedinger, 2000, 2001; Aleven et al., 2003).

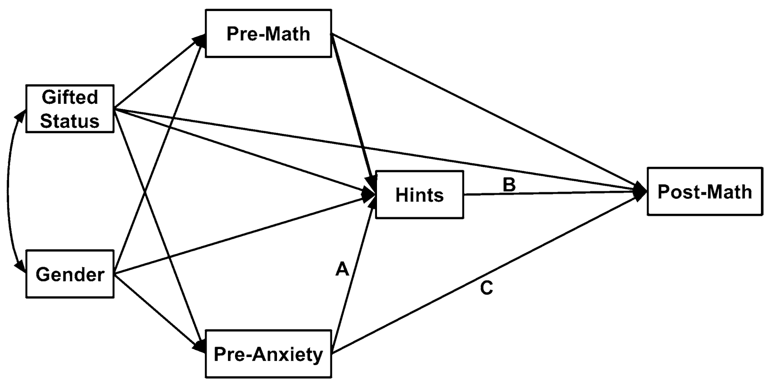

Fig. 1 The proposed multi-group structural equation model for the current study

3. The pattern of results (i.e., paths A and B) may be different between the problem set condition and the FH2T condition. Given the prior research on math anxiety and help-seeking behavior in online learning contexts, math anxiety may promote more inefficient help seeking (Aleven & Koedinger, 2000, 2001; Aleven et al., 2003). However, the hints in FH2T may be less anxiety inducing due to the game based nature and flexible problem solving that is less intimidating and does not penalize students for requesting help as is the case in the problem set condition. Therefore using hints in FH2T may not be as strongly related to math anxiety and math equivalence performance as in the problem set condition.

These hypotheses were pre registered and can be found at OSF https://doi.org/10.17605/OSF.IO/YWHXB. We do note that the original preregistration also included examining the roles of error feedback on the relation between math anxiety and performance of equivalence. However, due to the stark differences in how the two systems capture errors and error rates between the two conditions, we focused only on hint usage in the current study.

Methods

Participants

Ten teachers from six middle schools were recruited from a large, urban district in the Southeastern United States to participate in a 6-week randomized controlled trial. The 10 teachers together taught 29 math classes with a total of 689 students. Random assignment of the intervention condition occurred at the student-level, with the 689 students randomly assigned to FH2T ($n=348$) or problem set ($n=341$) conditions. Most students were in sixth grade (609 students, 88.4%), and the remaining (80 students, 11.6%) were in seventh grade. All students in the district were in one of three levels of math classrooms: advanced (525 students, 76.2%), on-level (111 students, 16.1%), or support (53 students, 7.7%). Most students in our sample were advanced sixth grade students. The racial breakdown of our initial sample is as follows: 48.4% was Asian, 39.0% was White, 6.4% was Hispanic, 2.6% was African American, and 3.6% was identified as other racial groups.

Due to scheduling constraints, 19 students from one classroom did not participate in the study, 195 students did not complete at least 50% of the items on the equivalence assessment, and nine students did not complete the math anxiety questionnaire at pretest. These 223 students were excluded in the following analyses. The 50% cutoff on pretest and posttest equivalence assessments was determined during preliminary analysis. The students who completed less than 50% of items tended to spend little time on the assessments (<13 out of 45 min), and dropped out. The 50% cutoff allowed us to more accurately estimate students' math understanding and learning. Excluded students had higher levels of pretest anxiety (excluded $M=18.1$, $SD=7.72$; included $M=15.7$, $SD=8.29$, $t(627)=-3.267$, $p=0.001$), were more likely to be in the seventh grade (26.1% of excluded, 4.5% of included, $\chi2(1, N=688)=69.148$, $p<0.001$) and White (50.8% of excluded, 34.1% of included, $\chi2(1, N=661)=16.013$, $p<0.001$), and less likely to be Asian (31.3% of excluded, 55.6% of included, $\chi2(1, N=661)=32.496$ $p<0.001$), or identified as gifted (29.2% of excluded vs 52.1% of included, $\chi2(1, N=661)=29.123$ $p<0.001$). The final sample included the remaining 466 students: 225 (48.3%) were in the FH2T condition, and 241 (51.7%) were in the problem set condition.

The racial breakdown of our final sample is as follows: 55.6% was Asian, 34.1% was White, 4.5% was Hispanic, 2.1% was African American, and 3.6% was identified as other racial groups. The student demographics and pretest scores of the final sample ($N = 466$) were comparable between conditions (e.g., 54.7% were male and 45.3% were female. Further, 52.1% of the students in our sample were identified as gifted in the data provided by the district (as indicated by the state criteria) and 47.9% were identified as not gifted (see Appendix for demographics based on condition). The study was approved by and conducted in accordance with the human subjects' guidelines of the Institutional Review Board.

Procedure

This study consisted of a 45-min pretest, four 30-min intervention sessions, and a 45-min posttest in a course of six weeks during the fall of 2019. During week 1, all students completed an assessment on their equivalence knowledge, and questionnaires on their math anxiety and self-efficacy. From weeks 2 to 5, students completed four intervention sessions on either FH2T or the problem set condition. The mathematical content was aligned between the two conditions. Finally, all students completed a posttest assessment along with questionnaires on math anxiety and self-efficacy in week 6. All study assignments were administered online in math classrooms during instructional periods, and students worked individually at their own pace using a device (see Chan et al., 2021 for details). For our analysis, we focused on the math anxiety questionnaire, pretest and posttest equivalence scores, and the use of hints in each condition.

Technology interventions

Two technology interventions were used in this study.

From Here to There!

In FH2T, students used gesture-actions to transform algebraic expressions (e.g., $11 + 55 + y + 89 + 45$) into mathematically equivalent goals (e.g., $100 + y + 100$). For instance, in Fig. 2, the student first dragged the 11 on top of 89 (Fig. 2b) to produce the sum of 100 on the right (addition; Fig. 2c). Then, the student performed a similar transformation with the 45 and 55 (drag to add; Fig. 2d) to produce 100 on the left (Fig. 2e). It is important to note that students could *attempt* mathematical errors (e.g., tapping the addition sign to combine $5 + 2x$) while transforming expressions in FH2T, but the system did not *commit* these errors. Rather, the system provided immediate feedback and allowed students to explore alternative moves (i.e., $5 + 2\times$shakes and remains as $5 + 2x$). When the active expression matched the goal state, a clover board appeared showing the number of clovers awarded based on the number of steps taken to reach the goal (Fig. 2f). Students received three clovers when they reached the goal in the fewest possible steps; they received two clovers when they reached the goal with one or two more than the fewest possible steps; they received one clover when they reached the goal with three or more than the fewest possible steps. The attempts of mathematical errors did not influence the number of clovers awarded.

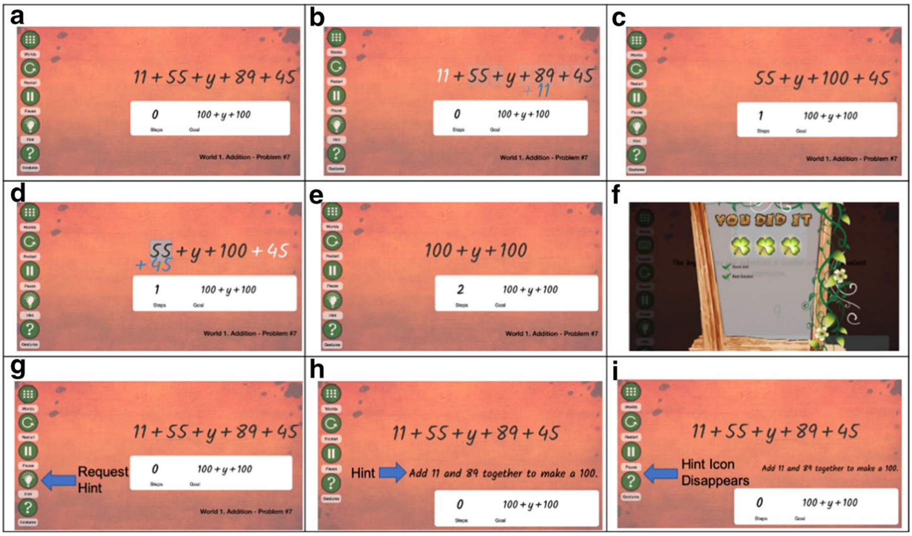

Fig. 2 A Sample Problem and Student Actions in From Here to There. *Note* Students transform a start expression (**a**), through a potential transformation process involving two steps (**b**–**d**) to reach the goal state (**e**). Students then will receive feedback (i.e., clovers) based on their performance (**f**). In each problem, one hint is available upon request. Students can request the hint by selecting the light bulb icon on the left side (**g**), and a hint is displayed underneath the problem (**h**). After requesting a hint, the hint icon disappears, and students cannot request additional hints (**i**)

Students worked at their own pace through 14 different levels of the game known as, "worlds" each containing 18 problems. Students needed to complete at least 14 problems in a world to move to the next world. In each problem in FH2T, students could request one hint that provided them with either a general strategy or a direct first-step instruction to successfully transform the starting expression into the goal state. For instance, in Fig. 2, when the student requested a hint (Fig. 2g), the hint stated "Add 11 and 89 together to make a 100" which provided an informative and strategic first step for students to reach the goal (Fig. 2h). Using hints in FH2T did not negatively impact the number of clovers a student could receive on that problem, and FH2T did not give any type of grade report at the end of each session. For each problem in FH2T, we recorded whether each student used the hint and coded "did not request a hint" as 0 and "did request a hint" as 1.

Problem set condition in ASSISTments

Students in the problem set condition solved traditional math problems in ASSISTments, a free online tutoring system for homework and problem-solving practice (Heffernan & Heffernan, 2014). The problems in ASSISTments were selected and adapted from three open-source middle-school math curricula: Utah Math Project (2016), Illustrative Mathematics (2017), and Engage NY (2014), so that the problems aligned with traditional instruction and the topics covered in FH2T. During each session, students received 24 to 39 questions one at a time, and they selected or entered their answer on the screen.

The problem set condition in ASSISTments include four types of problems: multiple choice, select all that apply, short answer and open response. Multiple choice questions involved selecting one answer by clicking the response option. Select all that apply involved selecting all response options that were correct for that question. Short answer involved typing the correct answer in a text box. Open response involved typing sentences into a text box for students to explain their thinking. All the problems had the options for requesting hints except for the open response questions which were ungraded and were used to gauge students' reasoning. The open response questions were excluded from the analysis.

For most problems in the problem set condition, students could request up to three hints, one at a time, by using the "request hint" button, on each problem (Fig. 3a). After a student asks for three hints, the button was then labeled "show answer" (Fig. 3b). Students could request the correct answer, enter the answer, and move on to the next problem (Fig. 3c). The amount of point students received depended on the number of incorrect answers they submitted and the number of hints they requested. Unlike FH2T, each use of a hint in the problem set condition involved a deduction of point for each problem. The point deduction varied depending on the type of problem and the number of answer options. For example, using a hint on a multiple-choice problem with only two options would result in more point deductions compared to using a hint on a short answer problem. Students received a green check if they answered the problem correctly without hints or incorrect attempts, and a red "X" if they ran out of points on the problem by using all hints or submitting multiple incorrect answers. We recorded the number of hints students requested and coded the data: "0" for "did not request any hints" and "1" for "did request at least one hint." This was done because the two technologies had a different number of hints available; FH2T had one hint per problem and the problem set condition had multiple hints.

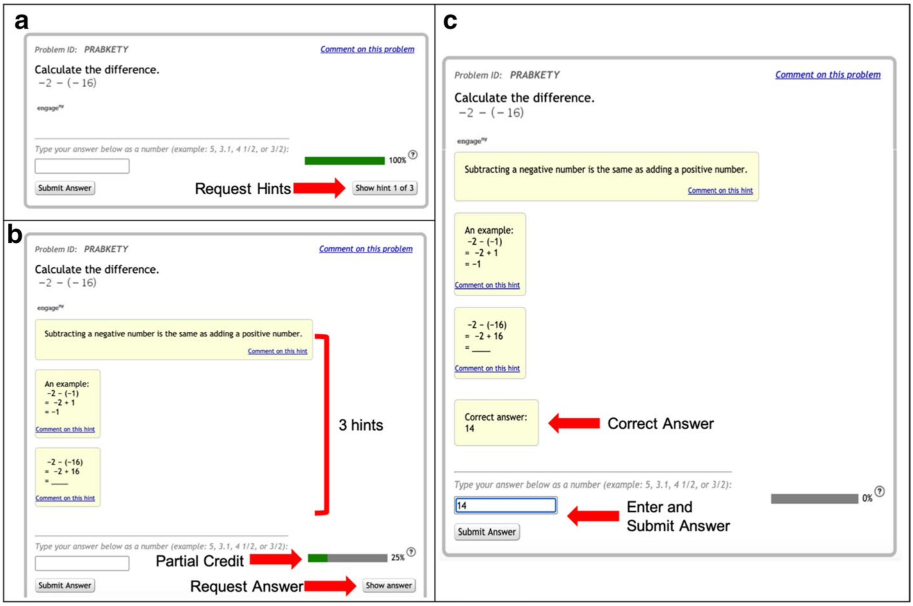

Fig. 3 A Sample Problem in the Problem Set Condition. *Note* **a** A sample problem in the problem set condition in ASSISTments with the hint request button available to students. **b** Each of the three hints is presented in a yellow box, and students can request the correct answer. **c** The correct answer is presented in a yellow box

Measures

All data was collected using the ASSISTments platform as students engaged in the study activites. A number of measures described below were collected as students solved problems.

Hint requests

In FH2T, there was one available hint for each problem. For each problem, we recorded whether students requested a hint (0) or not (1). We computed the percentage of problems on which each student requested hints (i.e., problems that a hint was requested / total problems completed) and used the percent hint request as a mediator in the analysis.

In the problem set condition in ASSISTments, most problems had three available hints. We recorded the number of hints students requested, and coded whether students requested any hints (1) or not (0) on each problem. We computed the percentage of problems on which a student requested hints (i.e., problems that at least one hint was requested / total problems completed) and used the percent hint request as a mediator in the analysis.

Pretest and posttest: equivalence understanding

Students' equivalence understanding was measured with six items selected from a previously validated measure of algebra knowledge (Rittle-Johnson et al., 2011; Star et al., 2015). An example of the item was identifying equivalent expressions (e.g., $(n+3)+(n+3)+(n+3)+(n+3)$ is equivalent to $4(n+3)$). Another example item was to recognize that if $10x+12=17$ was true, $10x+12-12=17-12$ would also be true. Each item was scored as correct (1) or incorrect (0), and the reliability of the items was fair, $KR\text{-}20=0.70$. The average score of the six items on the pretest was included as a covariate in the analyses. This measure is not intended to represent all aspects of math performance, but addresses specific aspects of math equivalence performance. (In this context, references to performance reflect math equivalence performance).

Math anxiety

Students' math anxiety was measured using 13 items selected from the Math Anxiety Scale for Young Children-Revised ($\alpha=0.87$; Ganley & McGraw, 2016). Although the Scale was developed for young children, it comprised items that measured students' negative reaction (4 items, e.g., "math gives me a stomachache"), numerical in-confidence (3 reverse items, e.g., "I like doing math problems on the board in front of the class"), and worrying (6 items, e.g. "I get worried before I take a math test"), providing opportunities for exploratory analyses on aspects of students' math anxiety. Further, because the items were designed for elementary school students, the variation in middle-schoolers' reading abilities would likely not interfere with students' self-report rating of their math anxiety. For each item, students rated how well it described their feeling towards math on a four-point scale (no = 0; not really = 1; kind of = 2; yes = 3). We reverse coded the numerical in-confidence items so higher scores reflected higher numerical in-confidence and higher math anxiety. The reliability of the scale ($\alpha=0.88$) was comparable to that reported in Ganley and McGraw (2016). While this measure was designed for younger students, similar reliability of this scale was found when used in a study of 8–13 year old students which found math anxiety to be a negative predictor of math fluency (Pollack et al., 2021). The average scores on the math anxiety scale at pretest was included as a predictor in the analyses.

Analytic approach

We examined the mediation of hint usage in both systems to test if there were differences in the relations between math anxiety, hint usage, and posttest math equivalence performance between the two conditions. We then created two SEM models, one for each technology, to examine how using hints differentially impacted the relation between pretest math anxiety and posttest math equivalence performance in the two conditions. In all models, we included students' gender and gifted status as covariates as they have been found to be associated with students' math anxiety and/or prior performance.

Results

Preliminary analyses

Descriptive statistics by condition are presented in Table 1 (further descriptive statistics such as participants' race, grade, and class are available in the Appendix). Reflecting the student-level random assignment, no significant between-groups differences were observed in any of the pre-intervention variables: gender ($\chi^2(1, N=466)=0.122$, $p=0.727$), gifted status ($\chi^2(1, N=466)=0.016$, $p=0.901$), pretest math performance ($t(464)=-0.720$, $p=0.472$), and pretest math anxiety ($t(464)=-0.317$, $p=0.752$). Between group differences were observed in hints ($t(464)=5.469$, $p<0.001$) such that there was greater use of hints in the problem set condition, and posttest math performance ($t(464)=-2.042$, $p=0.042$) where FH2T showed higher posttest math performance.

Table 2 reports the correlations among all the variables in the model, presented separately by condition. Although the size of correlations was often consistent across groups, several differences were observed, particularly among correlations involving the use of hints. For example, hint usage was strongly and negatively correlated with posttest equivalence scores in the problem set condition ($r=-0.581$); the same correlation was relatively small in the FH2T condition ($r=-0.160$). The same pattern was also observed between

Table 1 Descriptive statistics

	Problem set ($N=241$)				FH2T ($N=225$)			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Gender	0.46	0.50	0	1.00	0.44	0.50	0	1.00
Gifted	0.52	0.50	0	1.00	0.52	0.50	0	1.00
Pretest math anxiety	15.55	8.53	0	36.00	15.80	8.05	1.00	39.00
Hints	0.07	0.07	0	0.54	0.04	0.04	0	0.23
Pretest performance	3.76	1.63	0	6.00	3.87	1.59	0	6.00
Posttest performance	4.00	1.56	0	6.00	4.30	1.54	0	6.00

Gender: student gender (0 = male, 1 = female); Gifted: student gifted status (0 = not identified as gifted, 1 = identified as gifted)

Table 2 Correlations by intervention condition

Variable	Gender	Gifted status	Math anxiety	Percent hint	Pretest math	Posttest math
1. Gender		− 0.01 [− 0.14, 0.12]	0.20** [0.07, 0.32]	− 0.07 [− 0.20, 0.06]	− 0.06 [− 0.19, 0.07]	− 0.05 [− 0.18, 0.09]
2. Student Level	0.12 [− 0.00, 0.25]		− 0.20** [− 0.32, 0.07]	− 0.16* [− 0.29, 0.03]	0.48** [0.37, 0.57]	0.41** [0.30, 0.51]
3. Math anxiety	0.09 [− 0.04, 0.21]	− 0.15* [− 0.27, 0.03]		0.12 [− 0.01, 0.25]	− 0.41** [− 0.51, 0.29]	− 0.35** [− 0.46, − 0.23]
4. Hints	0.00 [− 0.13, 0.13]	− 0.41** [− 0.51, 0.30]	0.23** [0.11, 0.35]		− 0.21** [− 0.33, 0.08]	− 0.16* [− 0.28, − 0.03]
5. Pretest math	− 0.06 [− 0.19, 0.07]	0.51** [0.41, 0.60]	− 0.41** [− 0.51, 0.30]	− 0.50** [− 0.59, 0.40]		0.66** [0.57, 0.72]
6. Posttest math	− 0.10 [− 0.23, 0.02]	0.49** [0.38, 0.58]	− 0.33** [− 0.44, 0.21]	− 0.58** [− 0.66, 0.49]	0.64** [0.56, 0.71]	

Values in square brackets indicate the 95% confidence interval for each correlation. Top right diagonal is FH2T and bottom left diagonal is problem sets in ASSISTments

Gender student gender, *Gifted* student gifted status, *Math anxiety* pretest math anxiety, *Hints* percent hint usage, *Pretest math* pretest math equivalence performance, *Posttest math* posttest math equivalence performance.

*$p<0.05$, **$p<0.01$

hints and pretest performance (problem set: $r = -0.497$, FH2T: $r = -0.207$) and between gifted status and the use of hints (problem set: $r = -0.410$, FH2T: $r = -0.161$). In essence, these reflected a stronger association between the use of hints and other variables among students in the problem set condition, relative to those in the FH2T condition.

Primary analyses

Primary analyses involved a series of multi-group (FH2T vs. the problem set condition in ASSISTments) structural equation models (SEM), as illustrated in Fig. 1. Reflecting the power of the child-level randomization, an initial series of analyses found that all means (i.e., intercepts), paths, and errors that were solely based on the pre-randomization variables (female, gifted status, pretest math equivalence performance, pretest math anxiety) could be constrained equal without significantly reducing the model fit. Note that this involved constraining all paths that did not include hints or posttest math equivalence performance as equal. This served as the null model (Model 1, Table 3). The null model fit the data well, $\chi^2(16, N=466)=12.855$, $p=0.683$, CFI = 1.000, RMSEA = 0.000, SRMR = 0.025.

Eight follow-up analyses (Models 2–9, Table 3) tested whether each of the paths involving hints and/or posttest math equivalence performance could be constrained equally across groups, by comparing their fit indices to the null model. Results of those analyses found that four additional paths could be constrained equal across groups: The paths from gender to hints ($\chi^2(1, N=466)=0.635$, $p=0.426$), pretest math anxiety to hints ($\chi^2(1, N=466)=0.011$, $p=0.916$), gifted status to posttest math equivalence performance ($\chi^2(1, N=466)=0.166$, $p=0.684$), and pretest math anxiety to posttest math equivalence performance ($\chi^2(1, N=466)=0.070$, $p=0.791$). This resulted in a final model in which these additional four paths were also constrained to be equal across groups. As reported in Table 3 (Model 10), this final model fit reflected a more parsimonious fit than the null model ($\chi^2(4, N=466)=0.914, p=0.923$).

Path coefficients for the final model are presented in Table 4 and illustrated in Fig. 4. Note that the correlation between gifted status and gender was not statistically significant ($r=0.060$, $p=0.196$); however, the correlation between pretest math equivalence performance and pretest math anxiety was significant ($r=-0.370$, $p<0.001$), corroborating the previously documented negative relationship between achievement and math anxiety. Both were constrained equally across groups and are not shown in Fig. 4 in order to simplify the image. Also, for clarity, non-significant paths are presented in light gray.

As evidenced in Table 4 and Fig. 4, a clear group difference exists in terms of the role of hints. In the problem set condition, pretest math equivalence performance has a strong negative association with hints, which then has a strong negative association with posttest math performance—suggesting that part of the effect of pretest math equivalence performance on post math performance is being mediated by hint usage. However, these effects were not found in the FH2T condition. Collectively, this suggests that the relations of hint usage to prior performance and post performance may vary depending on the instructional design of the platform.

Discussion

The goals of the present study were to examine the relations between math anxiety, hint usage, and performance of math equivalence, and to consider the differences in these relations between two educational technologies: FH2T and a problem set condition with hints and immediate feedback. While there were no direct relations from math anxiety to hint usage, hints seemed to play a different role in each of these two technologies. Specifically, a significant negative relation from hint usage to later equivalence performance was found in the problem set condition, but not the FH2T condition. While these results do not provide a straightforward picture of the pathways through which math anxiety or hint-usage influences later math equivalence performance or determine which elements of the two technologies cause the different patterns of results, it suggests that variations in design components of educational technologies and students' initial performance can differentially influence learning.

Help-seeking behaviors are negatively related to math equivalence performance in problem set condition

The differences in the relations of hints and post-test performance in the problem set and the FH2T conditions encourage further exploration of the potential mechanisms that lead to more positive student outcomes. One plausible reason for the differences between the groups may be due to the ways that the technologies scaffold or encourage exploration of the math concepts or penalize students for using hints. Importantly, FH2T does not penalize a student when hints are requested or invalid mathematical actions are attempted, which may afford students opportunities for exploring different strategies to reach the goal state. Students may explore and acquire multiple strategies to the same goal through using hints, retrying problems, and comprehending immediate feedback on invalid actions. These feedback features in FH2T could support students by giving them the flexibility of applying various strategies and the opportunities to request hints and attempt invalid actions in a

Table 3 Summary of SEM models and model comparison

Model	Constrained paths	$\chi2$	df	p	$\chi2$ change	p change	CFI	RMSEA	SRMR	AIC
Null										
1	All pre-intervention effects	12.855	16	0.683			1.000	0.000	0.025	88.855
2	Gifted-Hints	17.830	17	0.400	4.975	0.026	0.999	0.010	0.034	91.830
3	Gender-Hints	13.490	17	0.703	0.635	0.426	1.000	0.000	0.027	87.796
4	PreMath-Hints	24.568	17	0.105	11.713	0.001	0.988	0.031	0.039	98.568
Additional constrained paths										
5	PreAnx-Hints	12.866	17	0.745	0.011	0.916	1.000	0.000	0.025	86.866
6	Gifted-PostMath	13.021	17	0.735	0.166	0.684	1.000	0.000	0.025	87.021
7	PreMath-PostMath	17.506	17	0.421	4.651	0.031	0.999	0.008	0.026	91.506
8	PreAnx-PostMath	12.925	17	0.741	0.070	0.791	1.000	0.000	0.025	86.925
9	Hints-PostMath	20.296	17	0.259	7.441	0.006	0.995	0.020	0.027	94.269
Final										
10	Constraints from 1,3,5,6,8	13.769	20	0.842	0.914	0.923	1.000	0.000	0.027	81.769

Models 2 through 10 compared to Model 1

Gifted student gifted status (0 = not identified as gifted, 1 = identified as gifted), *Hint* percent hint usage, *Gender* student gender (0 = male, 1 = female), *PreMath* pretest math performance, *PreAnx* pretest math anxiety, *PostMath* posttest math performance

Table 4 Final model: path coefficients

		Problem set condition			FH2T		
		β	C.R	*p*	β	C.R	*p*
Constrained Equal	Gifted-PreMath	0.501	12.475	< 0.001	0.501	12.475	< 0.001
	Gifted-PreAnx	– 0.182	– 4.013	< 0.001	– 0.182	– 4.013	< 0.001
	Female-PreMath	– 0.089	– 2.221	0.026	– 0.089	– 2.221	0.026
	Female-PreAnx	0.151	3.343	< 0.001	0.151	3.343	< 0.001
	Pre-Anx-PostMath	– 0.086	– 2.363	0.018	– 0.086	– 2.363	0.018
	Gifted-PostMath	0.141	3.654	< 0.001	0.141	3.654	< 0.001
	PreAnx-Hints	0.037	1.113	0.266	0.067	1.113	0.266
	Female-Hints	– 0.036	– 1.204	0.229	– 0.067	– 1.204	0.229
Unconstrained	PreMath-Hints	– 0.375	– 5.786	< 0.001	– 0.148	– 1.875	0.061
	Gifted-Hints	– 0.206	– 3.246	0.001	– 0.079	– 1.048	0.295

Gifted gifted status, *Hints* percent hint usage, *Female* student gender, *PreMath* pretest math equivalence performance, *PreAnx* pretest math anxiety, *PostMath* posttest math equivalence performance

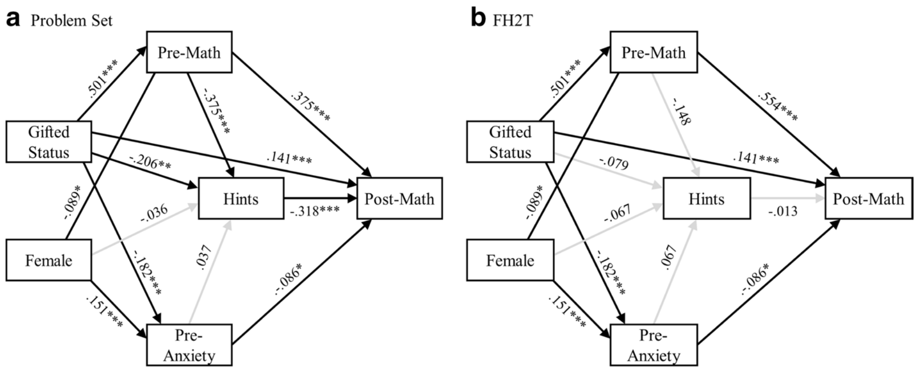

Fig. 4 Final SEM model

cognitively and emotionally safe environment. Alternatively, it could be that without a penalty for using hints, students regardless of their prior knowledge may feel more comfortable to check their understanding by using the hints, contributing to the weak association between pretest math equivalence performance and percent hint request in the FH2T condition.

In contrast, the problem set condition, as a traditional online tutoring system, has a different approach towards providing support for students. Because it is set up like a homework assignment or test evaluation, the system rewards students for entering the correct answer and penalizes students for requesting hints or submitting a wrong answer. Specifically, students receive correctness of feedback on each problem (e.g., a green check for a right answer or a red X for a wrong answer). Part of this feedback is also based on the number of their incorrect attempts and hint requests. For example, if they use all the available hints and request the correct answer, they will still receive a red X. Students also receive a score report on the problem set assignment at the end of each session showing their correctness on each problem. This evaluation-based system may create a more negative and intimidating environment for students when it comes to using hints because they are penalized on their score for using hints. Further, when students are not able to find the correct answer, they must use hints in order to advance to the next problem. While it is unclear if increased hint usage in the problem set condition is due to more unproductive help seeking behaviors, some students may be discouraged from using hints because it decreases their overall score on the assignment, whereas other students may abuse hints and the correct answer so that they can quickly complete the assignment. As suggested by prior research (Aleven & Koedinger, 2000, 2001; Aleven et al., 2003), both patterns of hint usage are ineffective and may hinder learning. Students, in general, request hints more often in ASSISTments than in FH2T because they cannot move on unless they enter the correct answer when solving problems. This may explain the negative link between hints and posttest scores in the problem set condition and not FH2T.

A third possibility for the differences in pathways may be that the innate design differences between the two technologies, independent from the hints. The game-based design of FH2T may have encouraged more flexible and exploratory thinking when solving the math problems more than the traditional problem sets in ASSISTments. As a game-based platform, FH2T supports exploration and encourages students to try many different approaches to solving a problem. Holistically, the game-based features or the dynamics of the FH2T system that support all mathematically valid actions and provide subtle immediate feedback on invalid actions may be a less intimidating environment for students to explore different solutions and may encourage them to use hints. For example, if students try to perform a step on the problem that is mathematically invalid, FH2T will provide immediate error feedback by sharking the expression and prevent them from committing this mistake. In the problem set condition, students are not prevented from making mistakes in the problem-solving process and may continue solving until they enter an incorrect answer. Thus, the differences in the ways that the two technologies respond to mistakes, may play a role in who requests a hint, and how requesting hints relates to their math performance. Future research will utilize log-data from student behaviors to explore the role of immediate feedback in the different systems.

Complex relations between math anxiety, help-seeking behavior, and performance

The results of this study demonstrate that the relations between prior knowledge, math anxiety, hint usage, and math outcomes are not straightforward. It is unclear whether the hints themselves or the students' math anxiety are the mechanisms that are leading to lower performance, or whether other factors, such as students' prior math knowledge and the design of the learning environments, influence the relations between hint usage, prior knowledge, and later performance of math equivalence. The negative association between hint usage and posttest equivalence may be partly driven by students' prior knowledge in math. In the problem set condition, lower knowledge students were more likely to request more hints, and requesting more hints was related to lower posttest performance of math equivalence. This contrasts previous work in ASSISTments which found that hints in the form of worked examples were effective for learning (Singh et al., 2011). Students with lower prior knowledge may need more help, which may lead them to request more hints. It is also plausible that greater hint usage may be a proxy for prior math knowledge. However, this account could not explain why the path between hints and posttest equivalence performance remained significant when accounting for the influences of pretest equivalence performance in the problem set condition but not the FH2T condition. This different pattern of results between the two conditions is especially noteworthy because the students showed comparable equivalence performance at pretest.

We also hypothesized that hint usage would mediate the association between math anxiety and later performance. As expected, math anxiety was negatively correlated with posttest scores on the equivalence knowledge assessment, with the effect comparable among students in the two technology interventions. Higher math anxiety was also correlated with more hints requests in the problem set condition, whereas this association was not significant in the FH2T condition. However, after accounting for students' initial math equivalence performance, this relation was considerably smaller..While this aligns with past findings that math anxiety may have negative relations to math learning and performance (e.g., Foley et al., 2017; Ma, 1999; Wu et al., 2012), it also reflects the close association and shared variance with concurrent math performance and highlights the importance of considering students' prior knowledge when examining the influences of math anxiety and hint usage on later performance. For example, it is plausible that students who have higher anxiety may be more likely to avoid math and therefore may not actively engage in the practice and problem solving. In a recent study, when students were asked to make decisions on the type of practice problems to engage in, the students with higher math anxiety chose easier problems that would be less likely to challenge them and increase their learning (Choe et al., 2019). Past research has also shown difficulty in teasing apart the phenomena of math anxiety from test anxiety. The anxiety of being assessed in math may not allow students to successfully demonstrate their knowledge (Kazelskis et al., 2000).

Limitations and future directions

To our knowledge this is one of the first studies that examines differences in pathways between hint usage and math anxiety in different types of instructional platforms. However, there are several important limitations that need to be considered. First, the sample consists of a high number of high performing students that are not representative of the United States a whole. In addition, we did see that some of the students who attrited were lower performing than those who stayed in the study. However, prior work has shown that there was no ceiling effect in the current sample and that both high and low performing students displayed learning gains using both of the two educational technologies (Chan et al., 2021). It may be beneficial for future research to stratify students based on performance level to further explore the association between prior knowledge and hint usage in these two educational technologies. Second, the math knowledge assessment consisted of only six items targeting students' understanding of equivalence, and did not comprehensively capture students' algebraic understanding (Kaput, 2008). However, equivalence is an important foundation and a strong proxy of students' algebraic understanding (Knuth et al., 2006). Future studies would benefit from examining the influences of math anxiety and hint usage on the broader algebraic knowledge as well as other topics of mathematics. Next, while we did find a significant effect of hint usage on posttest math equivalence performance in the problem set condition, additional assessments of both knowledge and hint usage at multiple time points would strengthen and extend the findings. Given students excluded for various reasons also had higher pretest anxiety levels, it is possible that the effect of anxiety on hint usage may be truncated; nevertheless, the significant correlation between pretest anxiety and posttest math equivalence performance suggests there is a meaningful and statistically useful degree of variability in the pretest anxiety scores in the final sample. Finally, only one conceptual model was considered in the present study; analyzing additional models may reveal other key variables. Despite these limitations, we believe that this study sheds light on the ways in which the fundamental design and affordances of educational technologies, such as the role of hint usage and feedback, influence learning. More experimental research is needed in order to understand the mechanisms and roles of hints (e.g., the number of hints given, correctness-focused hints vs process-based hints, consequences of requesting hints, etc.) on student performance.

Implications for instructional practice and technology design

Overall, this study demonstrates that there is a need for a better understanding of the role of hints in educational technologies and can inform the design of educational technologies that simultaneously promote math performance and productive help seeking behaviors in middle school students. Considering that we found no significant influence of math anxiety on hints and only a small effect of anxiety on math equivalence posttest equivalence performance in either technology condition, it is possible that the use of educational technologies—especially those with supportive hints and immediate feedback during problem practices—may ameliorate the previously found negative effects of math anxiety on performance. However, the different results on hint-usage and post-test scores between conditions provides implications about the specific design of hints within educational technologies. Designing hints to encourage exploration without being punitive may better support learning and prompt exploration of different problem-solving strategies regardless of their knowledge level. Correctness-based feedback or penalizing students for asking for help may not achieve the desired goals for providing students with the necessary supportive encouragement to scaffold learning. The design choices of how help and hints are presented in instruction and online educational technologies can differentially influence students' math behaviors and may encourage different approaches, leading to different pathways of learning. Teachers and instructional materials could provide students with more low-risk opportunities to ask for help and practice and display their knowledge in different ways, encouraging more productive help-seeking behaviors. These findings prompt thinking about how the fundamental design and affordances of educational technologies, such as the role of hint usage and feedback, influence learning. Further research is needed to examine how contexts of hints, such as the consequences of requesting hints, influence students' performance.

Appendix

Students' demographic information by condition, and their pretest scores.

	All ($N=466$)		FH2T ($n=225$)		Problem set ($n=241$)	
	n	%	*n*	%	*n*	%
Gender						
Male	255	54.7	125	55.6	130	53.9
Female	211	45.3	100	44.4	111	46.1
Race						
White	159	34.1	80	35.6	79	32.8
Asian	259	55.6	119	52.9	140	58.1
Hispanic	21	4.5	10	4.4	11	4.6
African American	10	2.1	6	2.7	4	1.7
Native American	5	1.1	1	0.4	4	1.7
Pacific Islander	1	0.2	1	0.4	0	0.0
Multi-racial	11	2.4	8	3.6	3	1.2
Grade						
Sixth	445	95.5	215	95.6	230	95.4
Seventh	21	4.5	10	4.4	11	4.6
Class						
Advanced	392	84.1	190	84.4	202	83.8
On-level	33	7.1	14	6.3	19	7.9
Support	41	8.8	21	9.3	20	8.3
Gifted status						
Gifted	243	52.1	118	52.4	125	51.9
Not gifted	223	47.9	107	47.6	116	48.1
Pretest math scores (*M*, *SD*)	3.82	1.61	3.87	1.59	3.76	1.63
Pretest math anxiety (*M*, *SD*)	15.67	8.29	15.80	8.05	15.55	8.53

Acknowledgements The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A180401 to Worcester Polytechnic Institute. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education. We would like to thank the teachers and students for their participation; the Summer Training in Arts and Sciences (STAR) Fellowship for funding our corresponding author for this research, David Brokaw and Graspable Inc. and Neil Heffernan and the ASSISTments Team for programming and data support.

Author contributions AI: Conceptualization, Writing, Project leader. EO: Conceptualization, Writing, Supervision, Project administration, Funding acquisition. VN: Writing, Project coordination. CM: External Independent Evaluator, Formal data analysis, Writing, Visualization. JY-CC: Conceptualization, Writing, Visualization, Project administration. HS: Writing, Data analysis. KCD: Writing. STS: Conceptualization, Writing.

Declarations

Competing interests Erin Ottmar was a designer and co-developer of From Here to There! The remaining authors have no competing interests to declare.

References

Aleven, V., & Koedinger, K. R. (2000). *Limitations of student control: Do students know when they need help? Lecture notes in computer science (Including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics*) (Vol. 1839, pp. 292–303). https://doi.org/10.1007/3-540-45108-0_33

Aleven, V., & Koedinger, K. R. (2001, May). Investigations into help seeking and learning with a cognitive tutor. *Papers of the AIED-2001 workshop on help provision and help seeking in interactive learning environments*.

Aleven, V., Elmar, S., Schworm, S., Fischer, F., & Wallace, R. (2003). Help seeking and help design in interactive learning environments. *Review of Educational Research, 73*(3), 277–320. https://doi.org/10.3102/00346543073003277

Anderson, J., Corbett, A., Koedinger, K., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of the Learning Sciences*, *4*(2), 167–207. https://www-jstor-org.ezpxy-web-p-u01.wpi.edu/stable/1466690?sid=primo&seq=4#metadata_info_tab_contents

Ashcraft, M. H. (2002). Math anxiety: Personal, educational, and cognitive consequences. *Current Directions in Psychological Science, 11*(5), 181–185. https://doi.org/10.1111/1467-8721.00196

Beilock, S. L., & Maloney, E. A. (2015). Math anxiety. *Policy Insights from the Behavioral and Brain Sciences, 2*(1), 4–12. https://doi.org/10.1177/2372732215601438

Chan, J. Y. C., Lee, J. E., Mason, C., Sawrey, K. B., & Ottmar, E. R. (2021). From Here to There!: A digital algebraic notation system impacts understanding of equivalence in middle school students. *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000596

Choe, K. W., Jenifer, J. B., Rozek, C. S., Berman, M. G., & Beilock, S. L. (2019). Calculated avoidance: Math anxiety predicts math avoidance in effort-based decision-making. *Science Advances, 5*(11), eaay1062. https://doi.org/10.1126/sciadv.aay1062

Chou, C.-Y., & Chang, C.-H. (2021). developing adaptive help-seeking regulation mechanisms for different help-seeking tendencies. *Educational Technology & Society, 24*(4), 54–66.

Foley, A. E., Herts, J. B., Borgonovi, F., Guerriero, S., Levine, S. C., & Beilock, S. L. (2017). The math anxiety-performance link. *Current Directions in Psychological Science, 26*(1), 52–58. https://doi.org/10.1177/0963721416672463

Ganley, C. M., & McGraw, A. L. (2016). The development and validation of a revised version of the math anxiety scale for young children. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2016.01181

Heffernan, N. T., & Heffernan, C. L. (2014). The ASSISTments ecosystem: Building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. *International Journal of Artificial Intelligence in Education, 24*(4), 470–497. https://doi.org/10.1007/s40593-014-0024-x

Hulse, T., Daigle, M., Manzo, D., Braith, L., Harrison, A., & Ottmar, E. (2019). From here to there! Elementary: A game-based approach to developing number sense and early algebraic understanding. *Educational Technology Research and Development, 67*(2), 423–441. https://doi.org/10.1007/s11423-019-09653-8

Hung, C.-M., Huang, I., & Hwang, G.-J. (2014). Effects of digital game-based learning on students' self-efficacy, motivation, anxiety, and achievements in learning mathematics. *Journal of Computers in Education, 1*(2–3), 151–166. https://doi.org/10.1007/s40692-014-0008-8

Inventado, P. S., Scupelli, P., Ostrow, K., Heffernan, N., Ocumpaugh, J., Almeda, V., & Slater, S. (2018). Contextual factors affecting hint utility. *International Journal of STEM Education, 5*(1), 1–13. https://doi.org/10.1186/s40594-018-0107-6

Iossi, L. (2007). Strategies for reducing math anxiety in post-secondary students. In S. M. Nielsen & M. S. Plakhotnik (Eds.), *Proceedings of the sixth annual college of education research conference: urban and international education section* (pp. 30–35).

Kaput, J. (2008). What is algebra? What is algebraic reasoning? In J. Kaput, D. Carraher, & M. Blanton (Eds.), *Algebra in the early grades* (pp. 5–17). Lawrence Erlbaum Associates.

Karabenick, S. A., & Berger, J.-L. (2013). Help seeking as a self-regulated learning strategy. In H. Bembenutty, T. J. Cleary, & A. Kitsantas (Eds.), *Applications of self-regulated learning across diverse disciplines: A tribute to Barry J. Zimmerman* (pp. 237–261). IAP Information Age Publishing.

Karabenick, S. A., & Knapp, J. R. (1988). Help seeking and the need for academic assistance. *Journal of Educational Psychology, 80*(3), 406–408. https://doi.org/10.1037/0022-0663.80.3.406

Karabenick, S. A., & Knapp, J. R. (1991). Relationship of academic help seeking to the use of learning strategies and other instrumental achievement behavior in college students. *Journal of Educational Psychology, 83*(2), 221–230. https://doi.org/10.1037/0022-0663.83.2.221

Kazelskis, R., Reeves, C., Kersh, M. E., Bailey, G., Cole, K., Larmon, M., et al. (2000). Mathematics anxiety and test anxiety: Separate constructs? *The Journal of Experimental Education, 68*(2), 137–146.

Kehrer, P., Kelly, K., & Heffernan, N. (2013). *Does immediate feedback while doing homework improve learning?* www.aaai.org

Kelly, K., Heffernan, N., Heffernan, C., Goldman, S., Pellegrino, J., & Soffer-Goldstein, D. (2013). Estimating the effect of web-based homework. *CEUR Workshop Proceedings, 1009*, 824–827. https://doi.org/10.1007/978-3-642-39112-5_122

Knuth, E. J., Stephens, A. C., McNeil, N. M., & Alibali, M. W. (2006). Does understanding the equal sign matter? Evidence from solving equations. *Journal for Research in Mathematics Education, 37*(4), 297–312.

Lyons, I. M., & Beilock, S. L. (2012). When math hurts: Math anxiety predicts pain network activation in anticipation of doing math. *PLoS ONE, 7*(10), e48076. https://doi.org/10.1371/journal.pone.0048076

Ma, X. (1999). A meta-analysis of the relationship between anxiety toward mathematics and achievement in mathematics. *Journal for Research in Mathematics Education, 30*(5), 520–540. https://doi.org/10.2307/749772

Maloney, E. A., Risko, E. F., Ansari, D., & Fugelsang, J. (2010). Mathematics anxiety affects counting but not subitizing during visual enumeration. *Cognition, 114*(2), 293–297. https://doi.org/10.1016/j.cognition.2009.09.013

Marwan, S., Dombe, A., & Price, T. W. (2020, June). Unproductive help-seeking in programming: What it is and how to address it. In *Proceedings of the 2020 ACM conference on innovation and technology in computer science education* (pp. 54–60). https://doi.org/10.1145/3341525.3387394

Newman, R. (2008). *The motivational role of adaptive help seeking in self-regulated learning.* In D. H. Schunk, & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications*. Lawrence Erlbaum Associates Publishers. https://psycnet.apa.org/record/2008-03967-013

Newman, R. S. (1990). Children's help-seeking in the classroom: The role of motivational factors and attitudes. *Journal of Educational Psychology, 82*, 71–80.

Ottmar, E. R., Landy, D., Goldstone, R. L., & Weitnauer, E. (2015a). Getting from here to there: Testing the effectiveness of an interactive mathematics intervention embedding perceptual learning. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings & P. P. Maglio (Eds.), *Proceedings of the thirty-seventh annual conference of the cognitive science society* (pp. 1793–1798). Pasadena, CA.

Ottmar, E., Landy, D., Weitnauer, E., & Goldstone, R. (2015b). Graspable mathematics: Using perceptual learning technology to discover algebraic notation. In M. Meletiou-Mavrotheris, K. Mavrou & E. Paparistodemou (Eds.), *Integrating touch-enabled and mobile devices into contemporary mathematics education* (pp. 24–48). Pennsylvania: IGI Global

Peeters, A., Robinson, V., & Rubie-Davies, C. (2020). Theories in use that explain adolescent help seeking and avoidance in mathematics. *Journal of Educational Psychology, 112*(3), 533–550. https://doi.org/10.1037/edu0000423

Pizzie, R. G., & Kraemer, D. J. M. (2017). Avoiding math on a rapid timescale: Emotional responsivity and anxious attention in math anxiety. *Brain and Cognition, 118*, 100–107. https://doi.org/10.1016/j.bandc.2017.08.004

Pollack, C., Wilmot, D., Centanni, T. M., Halverson, K., Frosch, I., D'Mello, A. M., et al. (2021). Anxiety, motivation, and competence in mathematics and reading for children with and without learning difficulties. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2021.704821

Price, T. W., Zhi, R., & Barnes, T. (2017). Hint generation under uncertainty: The effect of hint quality on help-seeking behavior. In E. André, R. Baker, X. Hu, M. Rodrigo, & B. Boulay (Eds.), *Artificial intelligence in education. AIED 2017. Lecture notes in computer science.* (Vol. 10331). Cham: Springer. https://doi.org/10.1007/978-3-319-61425-0_26

Ramirez, G., Shaw, S. T., & Maloney, E. A. (2018). Math anxiety: Past research, promising interventions, and a new interpretation framework. *Educational Psychologist, 53*(3), 145–164. https://doi.org/10.1080/00461520.2018.1447384

Razzaq, L., & Heffernan, N. T. (2010). Hints: Is it better to give or wait to be asked? *Lecture notes in computer science (Including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)*, *6094 LNCS*(PART 1) (pp. 349–358). https://doi.org/10.1007/978-3-642-13388-6_39

Renkl, A. (2002). Worked-out examples: instructional explanations support learning by self-explanations. In *Learning and instruction* (Vol. 12). www.elsevier.com/locate/learninstruc

Rittle-Johnson, B., Matthews, P. G., Taylor, R. S., & McEldoon, K. L. (2011). Assessing knowledge of mathematical equivalence: A construct-modeling approach. *Journal of Educational Psychology, 103*(1), 85–104. https://doi.org/10.1037/a0021334

Roschelle, J., Feng, M., Murphy, R. F., & Mason, C. A. (2016). Online mathematics homework increases student achievement. *AERA Open, 2*(4), 233285841667396. https://doi.org/10.1177/2332858416673968

Ryan, A. M., Gheen, M. H., & Midgley, C. (1998). Why do some students avoid asking for help? An examination of the interplay among students' academic efficacy, teachers' social-emotional role, and the classroom goal structure. *Journal of Educational Psychology, 90*(3), 528–535. https://doi.org/10.1037/0022-0663.90.3.528

Singh, R., Saleem, M., Pradhan, P., Heffernan, C., Heffernan, N. T., Razzaq, L., Dailey, M. D., O'Connor, C., & Mulcahy, C. (2011). Feedback during web-based homework: The role of hints. *Lecture notes in computer science (Including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)*, *6738 LNAI* (pp. 328–336). https://doi.org/10.1007/978-3-642-21869-9_43

Stamper, J., Barnes, T., & Croy, M. (2011). Enhancing the automatic generation of hints with expert seeding. *International Journal of Artificial Intelligence in Education, 21*, 153–167. https://doi.org/10.3233/JAI-2011-021

Star, J. R., Pollack, C., Durkin, K., Rittle-Johnson, B., Lynch, K., Newton, K., & Gogolen, C. (2015). Learning from comparison in algebra. *Contemporary Educational Psychology, 40*, 41–54. https://doi.org/10.1016/j.cedpsych.2014.05.005

Stowell, J., & Bennett, D. (2010). Effects of online testing on student exam performance and test anxiety. *Journal of Educational Computing Research, 42*(2), 161–171. https://doi.org/10.2190/EC.42.2.b

Szczygiel, M. (2020). Gender, general anxiety, math anxiety and math achievement in early school-age children. *Issues in Educational Research, 30*(3), 1126–1142.

Van Mier, H. I., Schleepen, T., & Van den Berg, F. (2019). Gender differences regarding the impact of math anxiety on arithmetic performance in second and fourth graders. *Frontiers in Psychology, 9*, 2690. https://doi.org/10.3389/fpsyg.2018.02690

Whitehill, J., & Seltzer, M. (2016). A crowdsourcing approach to collecting tutorial videos—Toward personalized learning-at-scale. In *L@S 2017 - Proceedings of the 4th (2017) ACM conference on learning at scale* (pp. 157–160). http://arxiv.org/abs/1606.09610

Wood, H., & Wood, D. (1999). Help seeking, learning and contingent tutoring. *Computers and Education, 33*(2–3), 153–169. https://doi.org/10.1016/S0360-1315(99)00030-5

Wu, S. S., Barth, M., Amin, H., Malcarne, V., & Menon, V. (2012). Math anxiety in second and third graders and its relation to mathematics achievement. *Frontiers in Psychology, 3*(JUN), 1–11. https://doi.org/10.3389/fpsyg.2012.00162