Educational Psychology Review (2026) 38:26
https://doi.org/10.1007/s10648-025-10097-1

INTERVENTION STUDY

The Powerful Teacher: A Power Hypothesis for the Benefits of Learning-by-Teaching

Sarah Shi Hui Wong[1,2] · Stephen Lim[3]

Received: 1 July 2025 / Accepted: 25 November 2025

Abstract
Why does teaching others enhance one's learning? This study advanced a novel power hypothesis of learning-by-teaching: When assuming the role of a teacher, students experience a heightened sense of power in influencing others, which increases their own learning. University students (N = 242) studied a scientific text using one of three learning methods: notetaking, explaining, or teaching. The notetaking group prepared to be tested and wrote study notes about the topic; the explaining group prepared to explain and wrote an explanation about the topic as how they would write a textbook for their peers; the teaching group prepared to teach and wrote a verbatim teaching script about the topic as how they would orate a lecture to their peers. All students were then tested on generating research questions that create new knowledge about the topic, as well as their basic comprehension. Teaching or explaining to others improved students' research question generation and comprehension more than writing study notes for their own learning. Crucially, teaching produced superior research question generation performance than explaining, and this advantage held even after controlling for comprehension. The teaching group reported experiencing more power than the explaining and notetaking groups, which mediated the advantage of teaching over explaining, as well as that of teaching and explaining over notetaking, for research question generation performance. An increased sense of power in a teacher potentially accounts for the learning benefits of teaching in enhancing one's research question generation performance. Indeed, teaching is powerful.

Keywords Learning by teaching · Explaining · Question generation · Power · Social influence

Extended author information available on the last page of the article

Published online: 27 February 2026

The Powerful Teacher: A Power Hypothesis for the Benefits of Learning-by-Teaching

"While we teach, we learn". Since this early observation by the Roman philosopher Seneca, decades of research has established that *learning-by-teaching*—taking on the role of a teacher and teaching others—is a powerful way to improve one's own learning of the material, relative to egocentric study activities (Allen & Feldman, 1973; Bargh & Schul, 1980; Duran & Topping, 2017; Fiorella & Mayer, 2013, 2014; Hoogerheide et al., 2014; Roscoe & Chi, 2007; for reviews, see Kobayashi, 2024; Lachner et al., 2022; Ribosa & Duran, 2022). The benefits of learning-by-teaching have been robustly observed across diverse outcomes including not only the tutor's basic recall, comprehension, and transfer (Lachner et al., 2022), but also more complex, higher order outcomes such as creative research question generation (Lim et al., 2024; Wong et al., 2023) and argumentative reasoning (Wong, 2025).

Learning-by-teaching has been viewed to comprise three distinct processes that contribute to the benefits of this technique: preparing to teach (i.e., studying the material with the expectation to teach; Benware & Deci, 1984; Fiorella & Mayer, 2014; Guerrero & Wiley, 2021; Kobayashi, 2024; Nestojko et al., 2014), actually teaching (Fiorella & Mayer, 2013, 2014), and responding to tutee or audience questions (Roscoe & Chi, 2008). Although learning-by-teaching has commonly been implemented via peer tutoring (e.g., Leung, 2019; Roscoe & Chi, 2007, 2008) or teachable agents in computer-based environments (e.g., Biswas et al., 2005, 2016; Chin et al., 2010), tutors can profit even when they do not interact with a real or remote audience in what has been termed *learning by non-interactive teaching* (Lachner et al., 2022), which the present study focuses on. For instance, students have been found to gain learning benefits from teaching a fictitious audience, whether by orally delivering audio- or video-recorded lectures (e.g., Fiorella & Mayer, 2013; Hoogerheide et al., 2014; Koh et al., 2018; Lachner et al., 2018; Wong et al., 2023) or writing verbatim teaching scripts as exact transcripts of how they would orate a lecture (i.e., *silent teaching*; Lim et al., 2021; Lim et al., 2024; Wong, 2025).

To date, however, there is no clear consensus on the underlying mechanisms that drive the benefits of learning-by-teaching. Whereas several theoretical accounts have been put forth, each of these accounts has also been contested (Lachner et al., 2022). Here, we propose that one promising new candidate mechanism is power—teaching others induces a heightened sense of power that improves the teacher's own learning. We first review extant theoretical accounts and evidence for and against each account, before discussing and testing the novel power hypothesis.

Extant Theoretical Accounts of Learning-by-Teaching

Three main non-mutually exclusive accounts have been proposed to explain the benefits of learning-by-teaching: (a) the retrieval hypothesis, (b) the generative hypothesis, and (c) the social presence hypothesis.

Retrieval Hypothesis

According to the retrieval hypothesis, learning-by-teaching benefits are attributable to retrieval practice (Koh et al., 2018). When students teach the material from memory, they engage in retrieval processes that have been well-established to promote durable learning (Karpicke, 2017; Roediger & Karpicke, 2006). Presumably, retrieving information from memory activates cue-relevant information through spreading activation and forms an elaborative semantic network that aids retention via multiple pathways to the target (Carpenter, 2009). Retrieval practice has also been theorized to involve reinstating and updating the episodic context associated with the retrieved information, such that the updated context representations serve as effective retrieval cues that aid retention (Karpicke et al., 2014).

In the first test of the retrieval hypothesis of learning-by-teaching, Koh et al. (2018) informed learners that they would later be asked to teach a text on the Doppler effect from memory, and that they should study and prepare to teach the text. After the preparation phase, learners were randomly assigned to complete one of four tasks: solving arithmetic problems (control group), taking a free recall test (retrieval practice group), or delivering a video-recorded lecture either unaided by any teaching materials (teaching-with-retrieval group) or by reading from a prescribed teaching script verbatim (teaching-no-retrieval group). On a delayed comprehension test one week later, both the teaching-with-retrieval and retrieval practice groups outperformed the control group, and did not differ in their performance. Crucially, teaching-with-retrieval and retrieval practice both outperformed teaching-no-retrieval, which did not differ from the control group. Based on these findings, Koh et al. (2018) concluded that learning-by-teaching improves learning only when it involves retrieving the material.

However, these interpretations in favor of the retrieval hypothesis have been challenged (Kobayashi, 2022; Lachner et al., 2022). Although Koh et al.'s (2018) teaching-no-retrieval group did not retrieve the material from memory, they also did not generate their own teaching explanations when reading the given teaching script aloud, unlike the teaching-with-retrieval group. To address this issue, Sibley et al. (2022) manipulated retrieval practice within non-interactive teaching by comparing closed- versus open-book teaching. In their study, learners read a text on the human respiratory system, then completed one of three tasks: restudying the text (control group), or delivering an audio-recorded oral explanation either with access to the text (open-book teaching group) or without access to the text (closed-book teaching group). On an immediate recall test, the open-book teaching group outperformed the closed-book teaching group, as mediated by richer oral explanations. Sibley et al. (2022) then concluded that teaching with retrieval is less effective when constrained by retrieval failure. However, these findings must be interpreted with caution as the study fundamentally failed to replicate the learning-by-teaching effect: Both the open- and closed-book teaching groups did not significantly differ from the restudy control on the recall test. As Sibley et al. (2022) noted, their use of an immediate test may also have stunted any benefits of retrieval practice, which tend to be larger on delayed tests (Rowland, 2014).

Generative Hypothesis

According to the generative hypothesis, learning-by-teaching evokes beneficial generative processes that improve learning (Fiorella & Mayer, 2016). During generative processing, students make sense of the to-be-learned material by selecting relevant information, organizing it into a coherent structure, and integrating it with their existing knowledge. Relatedly, when making sense of and teaching the material, students may engage in reflective knowledge-building where they metacognitively monitor and reflect on their understanding, and generate elaborations and inferences that go beyond the material to repair their knowledge gaps (Roscoe, 2014; Roscoe & Chi, 2007, 2008). Whereas the generative hypothesis does not specify the conditions under which teaching promotes generative processing, it is assumed that metacognition can be an antecedent in the orchestration of cognitive processes such as selecting, organizing, and integrating information (Fiorella & Mayer, 2015). Collectively, generative and metacognitive processes could thus enable students to build rich mental representations of the material that support their deep learning and knowledge generalization (Fiorella, 2023; Wittrock, 1974).

Relative to control methods such as self-explaining or writing study notes for one's own learning, learning-by-teaching has been found to increase students' generation of elaborations and metacognitive monitoring statements (Cheng et al., 2023; Lim et al., 2021, 2024; Wong, 2025), with some evidence for improved metacomprehension accuracy from teaching (Fukaya, 2013; Jacob et al., 2020; cf. Jacob et al., 2021; Lachner et al., 2020, 2021). Moreover, greater generative processing in producing more elaborations has been found to mediate the benefits of learning-by-teaching for test performance (Lachner et al., 2018; Wong, 2025).

However, there is limited evidence that increased metacognitive monitoring mediates learning-by-teaching benefits (Fiorella & Kuhlmann, 2020; Wong, 2025), even though it is an important component of reflective knowledge-building (Roscoe, 2014). Relatedly, although teaching others can produce better learning than sophisticated control methods such as retrieval practice (Jacob et al., 2020; Lachner et al., 2020; Wong et al., 2023; cf. Jacob et al., 2021; Lachner et al., 2021), it does not always trigger more generative processing or metacognitive monitoring than such control methods (Lachner et al., 2020, 2021).

Social Presence Hypothesis

According to the social presence hypothesis, learning-by-teaching induces feelings of social presence that could elicit (meta)cognitive and motivational processes that improve learning (Hoogerheide et al., 2016, 2019a). Social presence has been conceptualized as the extent that one views an audience—whether actual or imagined—as "real" and "present" (Kreijns et al., 2022; see also Gunawardena, 1995; Short et al., 1976), and has commonly been operationalized as the use of more self-other referential terms such as "you", "me", and "us" (e.g., Hoogerheide et al., 2016; Jacob et al., 2020; Lachner et al., 2018; Lim et al., 2021). Presumably, when students experience heightened social presence in teaching with an audience in mind, they are more likely to engage in beneficial adaptation processes (Clark & Brennan, 1991). For instance, students may anticipate what their audience knows or does not know (Nickerson, 1999) and adapt their teaching accordingly, such as generating more elaborations for less knowledgeable tutees (Wittwer et al., 2010). Moreover, higher social presence may elicit greater physiological arousal (Hoogerheide et al., 2019a) and enjoyment (Hoogerheide et al., 2019b) that boost the teacher's learning.

However, evidence for the social presence hypothesis is mixed. For instance, Jacob et al. (2020) reported initial evidence that oral explaining induces higher social presence than written explaining, thereby leading oral explainers to produce more comprehensive explanations that improved their inference test performance for a complex text. But in a later study when the same researchers induced higher social presence in written explaining by having students write their explanations in a chat messenger program instead of a text editor, both forms of written explaining produced comparable inference test performance as oral explaining and retrieval practice, even though the chat messenger group displayed higher social presence (Jacob et al., 2021; see also Lachner et al., 2018). Likewise, Wong (2025) found that writing teaching scripts rather than study notes increased students' social presence and performance on recall and argumentative reasoning tests, but that increased social presence was not associated with better test performance. Taken together, these findings suggest that increasing social presence in learning-by-teaching does not necessarily enhance learning outcomes (Herberg et al., 2012; Jacob et al., 2021; Ribosa & Duran, 2023; Wang et al., 2023).

A Power Hypothesis of Learning-by-Teaching

We advance a *power hypothesis* as a novel account of learning-by-teaching benefits—when taking on the role of a teacher and teaching others, students might experience a heightened sense of power that may be beneficial for their own learning. Whereas power has traditionally been defined as control over valued resources such as money or information (Galinsky et al., 2003; Keltner et al., 2003), power is also a psychological state that involves perceptions or subjective feelings of one's capacity to influence others (Anderson et al., 2012; Guinote, 2017; Overbeck & Park, 2001).

Accordingly, power is not simply a feeling of social presence—perceiving an actual or imagined audience as "real" and "present" (Kreijns et al., 2022) does not necessarily also imbue one with the sense of being able to steer and influence that audience. Power is also distinct from related constructs such as autonomy, self-efficacy, and productive agency. Whereas autonomy relates to one's capacity to volitionally steer one's own outcomes uncontrolled by others (Deci & Ryan, 2000; Ryan & Deci, 2000), power as influence is expressed in one's capacity to control others' outcomes, which could involve responsibility for others (Lammers et al., 2016). Likewise, whereas self-efficacy refers to the belief in one's own capabilities to execute the actions required for a particular outcome (i.e., feeling in control over tasks; Bandura, 1977), power encompasses an interpersonal form of control in regard to one's perceived ability to influence others (i.e., feeling in control over people; Korman et al., 2022). Finally, as developed in the context of collaborative learning, productive agency relates to the agency with which people choose to collaborate (or not) and produce in their environment, as expressed in the effort and motivation to construct shared meaning with others (Schwartz, 1999; Schwartz & Lin, 2001). For instance, in learning-by-teaching, observing one's tutees use what they have been taught provides recursive feedback that affirms one's productive agency in producing and contributing one's ideas and having them realized in the world (Okita & Schwartz, 2013). Thus, productive agency does not necessarily implicate an interpersonal power dynamic or differential, instead focusing on one's capacity to express agency in production and original contribution. In contrast, a sense of power relates specifically to one's capacity to influence others, and can theoretically manifest even in non-interactive teaching when students do not interact with their audience or witness how their teaching is received.

According to French and Raven's (1959) now-classic typology of power bases, power can stem from various sources such as the ability to reward or punish, or the target's identification with the influencing agent, or one's legitimate right to influence (e.g., based on social structures involving a hierarchy of authority), or one's expertise or superior knowledge. Applied in the classroom, a teacher may hold power based on their students' expectations that they will be rewarded or punished for complying with the teacher's influence attempt or not, or attraction to and identification with the teacher, or perceptions of the teacher's legitimate right to make particular requests in their position as "teacher", or perceptions of the teacher as competent and knowledgeable in particular areas (McCroskey & Richmond, 1983). Thus, when students step into the role of a teacher in learning-by-teaching, they may experience a greater sense of power that is associated with this role. In line with this idea, giving advice—regardless of whether it is solicited or not—has been found to enhance the advisor's sense of having exercised influence over others' behavior and outcomes, thereby instilling the advisor with a sense of power (Schaerer et al., 2018).

In turn, when students experience heightened power as teachers, they may enact behaviors that they perceive as defined by this role. Even as perceived by undergraduates and academics in higher education today, the traditional socially structured student–teacher power relationship is characterized as one constituted through a teacher's authoritative knowledge and a student's deference to that authority, as well as the teacher's power to affirm the student and build self-esteem (Symonds, 2021). Indeed, high status is often associated with competence (Fiske et al., 2007). Thoroughly mastering the material to be taught is a critical requirement that is intrinsic to the role of teacher, such that teachers who do not know the material well face the prospect of embarrassment before their students (Allen & Feldman, 1973), who view not knowing the subject matter as a type of "teacher misbehavior" (Kearney et al., 1991). Accordingly, in learning-by-teaching, power in the role of a teacher could promote a sense of responsibility (Okita & Schwartz, 2013) that drives students to dutifully improve their own learning of the material and teach their audience well.

Indeed, power triggers a generalized approach orientation, energizes thought, speech, and action, and exerts cognitive and motivational effects that improve task performance (Galinsky et al., 2003; Guinote, 2017; Keltner et al., 2003). When people are placed in a position of power, they experience cognitive benefits in attending to key information, inhibiting peripheral information, and flexibly varying their attentional focus depending on task demands (Guinote, 2007a). Conversely, lacking power impairs executive functions in updating, inhibiting, and planning (Smith et al., 2008). Power also motivates self-regulation toward effective performance (DeWall et al., 2011) by increasing goal-directed action, persistence, and flexibility that facilitate goal pursuit (Guinote, 2007b). This motivating force of power resonates with findings from learning-by-teaching research that studying with the expectation to teach (i.e., taking on the role of a teacher) increases students' intrinsic motivation and learning, relative to studying with the expectation to be tested (i.e., taking on the role of a learner; Benware & Deci, 1984). Likewise, when people take on the role of an advisor to give others advice about a goal that they themselves are struggling with, they experience a boost in confidence that increases their own motivation to pursue that goal (Eskreis-Winkler et al., 2018).

The Present Study

Altogether, extant theoretical and empirical literature provides the bases for our novel power hypothesis of learning-by-teaching. Specifically, when students take on the role of a teacher and teach others, they might experience a heightened sense of power that may be beneficial for their own deep learning. This study tested the power hypothesis for the benefits of learning-by-teaching, with a focus on students' research question generation performance.

Developing students' proficiency in asking research questions is an educationally valued goal that is key to scientific literacy, as set forth in the National Research Council's (2013) *Next Generation Science Standards*. Indeed, asking good research questions forms a cornerstone of scientific inquiry and discovery when students construct knowledge and solve problems (Pedaste et al., 2015). Posing high-quality questions catalyzes further inquiry processes in determining what information should be sought (Kedrick et al., 2023; Tawfik et al., 2020) and reveals the depth of students' thinking and learning (Chin & Brown, 2002). As classified in questioning hierarchies, good research questions align with the peak *create* level of Bloom's taxonomy (Anderson et al., 2001; Bloom, 1956), in that such questions go beyond existing knowledge to create new knowledge (Chin & Brown, 2002; Dillon, 1984; Tawfik et al., 2020). Whereas less sophisticated questions tend to be factual with readily available answers or simply describe or compare various phenomena, *create* questions align with expert-like reasoning in combining or reorganizing different elements to construct new knowledge or hypotheses, thereby driving innovation (Tawfik et al., 2020; Wong et al., 2023).

Prior studies have demonstrated that learning-by-teaching enhances *create* question generation more than control learning methods such as writing study notes (Lim et al., 2024) or even retrieval practice and concept-mapping (Wong et al., 2023). However, the mechanisms underlying the teaching advantage remain unresolved. In the aforementioned studies, participants taught with access to the to-be-learned material such that there was no need or grounds for them to retrieve it from memory, thus rendering the retrieval hypothesis (Koh et al., 2018) inapplicable to the observed learning benefits of teaching. Still, based on the social presence hypothesis (Hoogerheide et al., 2016) and generative hypothesis (Fiorella & Mayer, 2016), it is possible that teaching others induced higher levels of social presence and/or generative processing that improved participants' learning. Hence, we concurrently examined these hypotheses alongside the power hypothesis.

University students were first trained on question generation, then randomly assigned to study a text using one of three learning methods: notetaking, explaining, or teaching. In the notetaking control condition, participants prepared to be tested by writing study notes about the text, then refined and revised their notes. In the explaining condition, participants prepared to explain the text by writing explanatory notes about it, then wrote an explanation on the text as how they would personally word it when writing their own textbook on the topic for their peers. In the teaching condition, participants prepared to teach the text by writing teaching notes about it, then wrote a verbatim (i.e., word-for-word) teaching script on the text as how they would exactly orally deliver an actual lecture on the topic to their peers (Lim et al., 2021, 2024; Wong, 2025). All participants then rated their sense of power and learning experience during the intervention, and predicted their test performance. Finally, participants were tested on generating as many *create* questions as possible about the text, as well as their basic comprehension of the text.

Holding constant the modality of learning in written format across all three conditions enabled us to distill the unique effects of "teaching", particularly in view of findings that oral versus written teaching may exert differential effects on learning (Lachner et al., 2022; cf. Lim et al., 2021). During the intervention, all participants also wrote their responses under open-book conditions such that there was no need or grounds for them to retrieve the material from memory, thus foreclosing the retrieval hypothesis (Koh et al., 2018) from accounting for any learning benefits of teaching. In addition, ensuring that all participants studied the material with an expectancy—either that they would later be tested on or be asked to explain or teach the content—meant that the presence of learning goals (or lack thereof) per se could not account for any subsequent differences in test performance across conditions.

Indeed, meta-analytic evidence (Kobayashi, 2024) has found that: Relative to merely studying without teaching expectancy (i.e., preparing to teach), teaching after studying *with* teaching expectancy yields a medium learning benefit ($g=0.48$), whereas the effect of teaching after studying *without* teaching expectancy does not differ from zero ($g=-0.02$). Moreover, actually teaching after preparing to teach yields a small-to-medium benefit ($g=0.38$) over preparing to teach only. Thus, preparing to teach and actually teaching both uniquely contribute to learning-by-teaching effects. Accordingly, the teaching and explaining conditions in this study both included an expectancy (preparing to teach vs. explain) and actually communicating knowledge in written modality to the same fictitious target audience (participants' peers) with the intention of helping them learn (actually teaching vs. explaining). Crucially, although the teaching and explaining conditions both involved generating materials for others' learning (lecture transcripts or textbook explanations for one's peers; Ribosa & Duran, 2022), the key difference was that the teaching condition wielded the intentionality of taking on the role of a "teacher" (Kobayashi, 2023). Indeed, the intention to teach matters in characterizing the act of "teaching"—whereas all teachers explain the to-be-learned material, not all who explain are teachers.

Concurrently, we examined the social presence hypothesis and generative hypothesis by scoring students' written study notes, textbook explanations, and teaching scripts on the number of: (a) self-other referential terms such as "me" and "you," which served as a proxy for perceived social presence (e.g., Hoogerheide et al., 2016; Jacob et al., 2020; Lachner et al., 2018; Lim et al., 2021), (b) elaborations (e.g., Jacob et al., 2020; Lachner et al., 2018), and (c) monitoring statements (e.g., Roscoe, 2014). According to the social presence hypothesis, learning-by-teaching would trigger the use of more self-other referential terms, in turn mediating any advantage of teaching on test performance; according to the generative hypothesis, learning-by-teaching would stimulate generative and metacognitive processing (i.e., greater reflective knowledge-building) in the form of more elaborations and monitoring statements, in turn mediating any advantage of teaching on test performance. Hence, we examined: (a) whether the three learning groups significantly differed in their use of self-other referential terms, elaborations, and monitoring statements, (b) whether these process measures correlated with participants' *create* question generation performance, and if so, (c) whether these process measures mediated the effects of learning method.

Based on extant evidence for the benefits of generating materials for others' learning (Ribosa & Duran, 2022) and, specifically, a learning-by-teaching benefit in *create* question generation (Lim et al., 2024; Wong et al., 2023), we predicted that the teaching and explaining groups would outperform the notetaking group in generating more *create* questions about the studied material, with a greater boost from teaching than explaining. Importantly, based on the power hypothesis: To the extent that students experience a greater sense of power when taking on the role of a "teacher" to write teaching scripts than when writing textbook explanations or study notes for their own learning, this heightened power would mediate the advantage of teaching over explaining and notetaking for students' *create* question generation performance. Additionally, based on the social presence hypothesis and generative hypothesis, respectively, one would expect teaching to induce greater perceived social presence as evidenced by the use of more self-other referential terms, as well as greater generative and metacognitive processing as evidenced by the use of more elaborations and monitoring statements, in turn mediating any benefits of teaching on *create* question generation performance.

Method

Participants and Design

The participants were 242 students (161 women, 78 men, 3 undisclosed) between the ages of 18 and 47 ($M=21.55$, $SD=3.42$) from a large public university in Singapore, who received either course credit or monetary reimbursement for their participation. The students came from a range of majors and years in college. Based on the weighted mean effect size of non-interactive teaching with preparing-to-teach over a non-teaching control ($g=0.48$) in Kobayashi's (2019) meta-analysis, a power analysis (G*Power; Faul et al., 2007) indicated that at least 70 participants per condition were needed to observe a learning-by-teaching effect for two-tailed between-subjects pairwise comparisons (i.e., t tests) at 80% power and $\alpha=.05$.

This study used a between-subjects design with learning method as the key independent variable, whereby participants were randomly assigned to one of three conditions: notetaking (control group; $n=81$), explaining ($n=80$), or teaching ($n=81$). To ascertain that any effects of the learning methods generalized across topics, study text was included as a second independent variable for control purposes, whereby participants were randomly assigned to study a text on either "origins of the moon" or "vanishing songbirds". The main learning outcomes of interest were: (a) participants' question generation test performance, as assessed by the number of *create*-level questions that they produced, and (b) participants' comprehension test performance, as assessed by the number of short-answer comprehension questions that they correctly answered.

All participants provided their written informed consent. This study was conducted with ethics approval from the university's institutional review board.

Materials and Procedure

A flowchart of the experimental procedure is depicted in Fig. 1.

Question Generation Training

Upon arriving for the experiment, all participants were first trained on question generation based on training protocols from Wong et al. (2023), so that they understood what was required of them in generating *create* questions. To ensure participants' understanding, the experimenter verbally reinforced all training and task instructions during the experiment. Participants received a training handout (see Table 1) that introduced and explained the features of all six question levels corresponding to Bloom's taxonomy, and that included sample action prompts associated with each level: *remember*, *understand*, *apply*, *analyze*, *evaluate*, and *create*. Participants were instructed on all question levels to guide their holistic understanding of the various question types and their differentiation of questions that fulfilled the *create* level versus those that did not (i.e., non-*create* questions).

Next, participants were presented with a brief 199-word practice text on "enzymes" (adapted from Meyer, 1975; Wong et al., 2023; available in the online supplemental materials), which did not relate to either of the critical study texts. Participants were given 5 min to practice generating at least two *create* questions based on the "enzymes" text with reference to the question levels training handout. After 5 min, all participants received concise verbal feedback on the questions that they had generated. Specifically, they were advised within a couple of sentences whether their questions correctly reflected the *create* level, and if not, how they could modify their questions appropriately (e.g., Wong et al., 2023).

Pre-Studying Questionnaire

Next, participants completed an 8-item pre-studying questionnaire that measured their perceived prior knowledge of the topic, predicted test performance, and interest in the topic. All ratings were made on a 5-point scale. Specifically, participants indi-

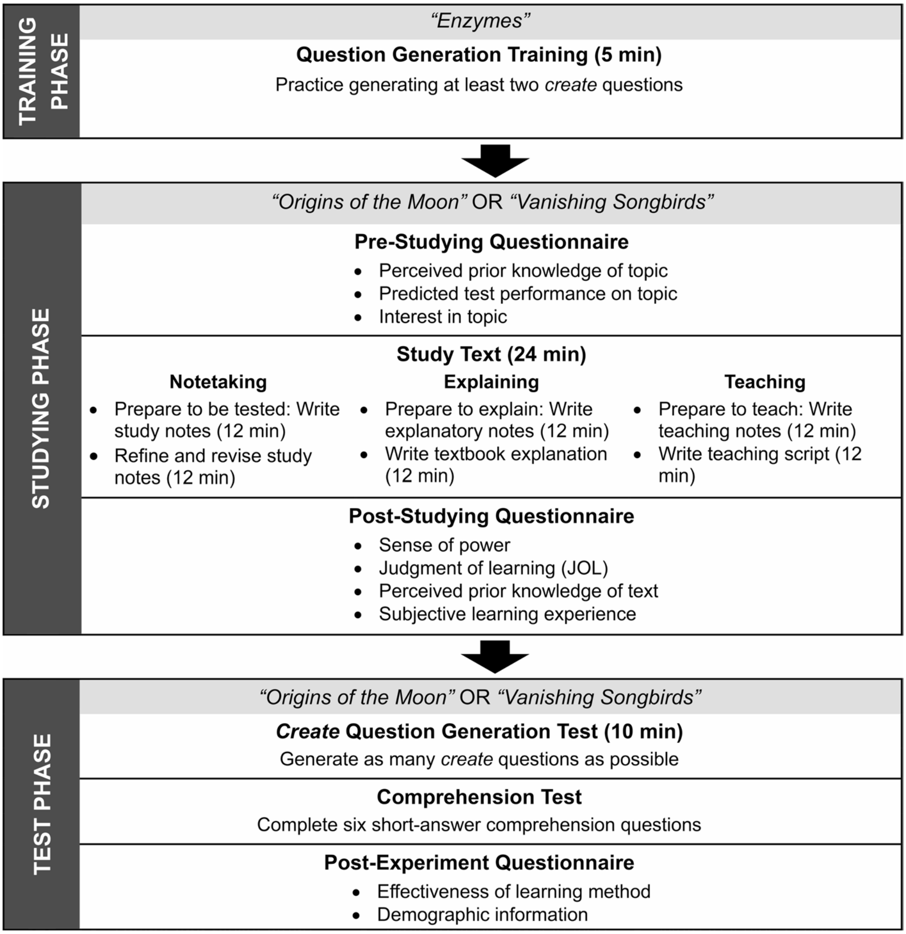

Fig. 1 Flowchart of experimental procedure

cated their perceived prior knowledge of the to-be-learned topic: "How much prior knowledge do you have about the (origins of the moon/vanishing songbird population)?" (1 = *not at all*; 5 = *a lot*). They also predicted their test performance on the topic: "How well do you think you will do on a test on the (origins of the moon/vanishing songbird population)?" (1 = *very poorly*; 5 = *very well*). Further, participants rated their topic interest (1 = *strongly disagree*; 5 = *strongly agree*) on six items adapted from Clinton and van den Broek (2012): (a) "I am interested in reading about (the moon/songbirds).", (b) "Reading about (the moon/songbirds) is boring for me." (reverse-scored), (c) "I am engaged when reading about (the moon/songbirds).", (d) "I am interested in reading about (astronomy/forestry).", (e) "Reading about (astronomy/forestry) is boring for me." (reverse-scored), and (f) "I am engaged when reading about (astronomy/forestry).". Participants' mean topic interest was computed as their mean rating across the six items. The topic interest measure demonstrated good internal consistency in this study, Cronbach's $\alpha = .89$.

Table 1 Question generation training handout: Question levels based on Bloom's taxonomy

Level	Category	Associated Cognitive Processes	Sample Action Prompts
1	Remember	Answer requires *recall/remembering* of terminology, specific facts, definitions, and basic concepts covered in the text	Identify, recognize, indicate, list, name specific events, locations, people, dates, sources of information (e.g., Who? What? Where? When? Which?)
2	Understand	Answer requires *basic understanding* (i.e., descriptions, explanations, examples) of concepts in the text	Describe, explain, give examples of, summarize, generalize
3	Apply	Answer requires *using/applying* acquired knowledge, facts, and concepts in a new situation or in a different way	Predict, give other examples in other contexts, seek exceptions
4	Analyze	Answer requires *examining and breaking down information* into constituent parts by identifying motives/causes, making inferences and finding evidence to support generalizations, or seeking causes and/or consequences	Compare, contrast, differentiate, organize, deconstruct
5	Evaluate	Answer requires *making judgments* about information, validity of ideas, or quality of work based on a set of criteria	Appraise, assess how effective/optimal or which is most important/valuable, check for discrepancies/inconsistencies in information
6	Create	Answer requires *creating new knowledge, ideas, or perspectives* by compiling information in a different way, combining elements in a new pattern, or proposing alternative solutions	Adapt, produce alternative hypotheses or solutions

Reprinted from "To Ask Better Questions, Teach: Learning-by-Teaching Enhances Research Question Generation More Than Retrieval Practice and Concept-Mapping" by S. S. H. Wong, K. Y. L. Lim, and S. W. H. Lim, 2023, *Journal of Educational Psychology*, *115*(6), p. 800. Copyright 2023 by American Psychological Association.

Study Texts and Learning Methods

Participants then began the intervention, during which they studied a text on either "origins of the moon" or "vanishing songbirds" (adapted from Clinton & van den Broek, 2012; available in the online supplemental materials). The texts had Flesch-Kincaid grade levels of 12.3 and 13.7, respectively. Each text comprised 730 words presented in four paragraphs, and discussed hypotheses for their respective scientific phenomena. All participants were given 24 min to study their text using their randomly assigned learning method. Thus, the study duration was exactly matched across all three learning conditions.

In the *notetaking* control condition, participants were told that they would later be tested on the text content. The nature of the test was not disclosed. To prepare for the test, participants were told that they would be studying the text and writing their own study notes about the topic, and that they would later be refining and revising their notes. Participants were given 12 min to write their study notes with reference to the study text. Then, the text was collected back and participants were given 12 min to refine and revise their notes.

In the *explaining* condition, participants were told that they would later be asked to explain the text content. The nature of explaining was not disclosed. To prepare to explain, participants were told that they would be studying the text and writing their own explanatory notes about the topic, and that they would later be explaining based on their notes. Participants were given 12 min to write their explanatory notes with reference to the study text. Then, the text was collected back and participants were given 12 min to write an explanation on the text, as how they would personally word it when writing their own textbook on the topic for their peers.

In the *teaching* condition, participants were told that they would later be asked to teach the text content. The nature of teaching was not disclosed. To prepare to teach, participants were told that they would be studying the text and writing their own teaching notes about the topic, and that they would later be teaching based on their notes. Participants were given 12 min to write their teaching notes with reference to the study text. Then, the text was collected back and participants were given 12 min to write a verbatim (i.e., word-for-word) teaching script on the text, as how they would exactly orally deliver an actual lecture about the topic to their peers (Lim et al., 2021, 2024). A brief lecture opening was provided as an example of a verbatim teaching script: "Good day, everyone! In this lecture, we will learn about ("origins of the moon"/"vanishing songbirds"). Let's begin…".

Across all three conditions, participants completed the studying task independently without feedback on their responses, and were told that their study notes or textbook explanation or teaching script would later be viewed by an audience for educational or research purposes. Because the learning benefits of teaching have been attributed, at least in part, to retrieval practice when teaching unaided from memory (Koh et al., 2018), participants in the explaining and teaching conditions were allowed to refer to their notes when writing their textbook explanation or teaching script, just as how notetaking participants could access their study notes when refining them. This procedure dissociated teaching from retrieval-based learning, ensuring that any observed learning benefits of teaching in this study could not be due to retrieval practice.

Post-Studying Questionnaire

After the 24-min studying period, participants completed a post-studying questionnaire in which they reported their sense of power and learning experience during the intervention, and predicted their test performance.

Sense of Power To measure participants' sense of power—perceptions or subjective feelings of one's capacity to influence others—during the intervention, we extracted four items from extant established measures: Anderson et al.'s (2012) Personal Sense of Power Scale, the Dominance factor of Cassidy and Lynn's (1989) Achievement Motivation Scale, and Schaerer et al.'s (2018) perceived influence measure. We adapted these items for the present learning-by-teaching context, where each item began with "When I was writing (study notes/an explanation/a teaching script as a teacher),": (a) "I can get others to listen to me.", (b) "I have authority over other people.", (c) "I enjoy planning and deciding how other people should learn.", and (d) "I feel I have the ability to influence other people.". All items were rated on a scale from 1 (*strongly disagree*) to 7 (*strongly agree*). Participants' mean sense of power was computed as their mean rating across all four items. The sense of power measure demonstrated high internal consistency in this study, Cronbach's α = .83.

Judgment of Learning (JOL) Participants then made a post-studying metacognitive JOL in predicting how much of the material from the study text they would remember later on an 11-point scale (0%, 10%, 20%, …, 100%).

Perceived Prior Knowledge of Text Participants reported how well they knew the subject matter in the study text prior to reading it (1 = *not very well*; 7 = *very well*).

Subjective Learning Experience Participants also rated their subjective learning experience during the intervention. First, they indicated how understandable the study text was (1 = *not at all*; 7 = *extremely*). Then, they rated eight phenomenological items (adapted from Fiorella & Mayer, 2013, 2014) on a scale from 1 (*strongly disagree*) to 7 (*strongly agree*) in terms of: (a) the difficulty of the subject matter ("*difficult*"), (b) how much they enjoyed learning about the subject matter ("*enjoyment*"), (c) how much they would like to learn with the learning method in the future ("*future*"), (d) their perceived understanding of the subject matter ("*understand*"), (e) their interest in learning more about the subject matter ("*interest*"), (f) how useful they found the subject matter to them ("*useful*"), and (g) how stressed they felt while learning about the subject matter ("*stress*"). They also rated the amount of mental effort they had invested while learning the subject matter ("*effort*") on a scale from 1 (*very low effort*) to 7 (*very high effort*).

Create *Question Generation Test*

After completing the post-studying questionnaire, participants began the question generation test in which they were given 10 min to generate and write down as many *create*-level questions as possible about the text they had studied (either "origins of the moon" or "vanishing songbirds"). Participants were reminded to generate only *create*-level questions, rather than questions from any other level of Bloom's taxonomy.

Comprehension Test

Then, participants completed a self-paced comprehension test on the text they had studied (either "origins of the moon" or "vanishing songbirds"). The comprehension test comprised six short-answer questions (adapted from Clinton & van den Broek, 2012; available in the online supplemental materials) that required participants to recall information from the text, or to connect information explicitly stated in the text, or to apply information from the text. A sample comprehension test question for the "origins of the moon" text was: "Describe the 'double planet hypothesis' of the origin of the moon. State one scientific fact or research finding supporting that particular theory.". A sample test question for the "vanishing songbirds" text was: "How do nest parasites affect neotropical songbird populations? Include one research finding in your answer.".

Post-Experiment Questionnaire

Finally, after completing the tests, participants responded to a post-experiment questionnaire in which they rated the effectiveness of their learning method in helping them learn the material (1 = *not at all*; 7 = *extremely*), and provided their demographic information. Participants were then debriefed and thanked.

Scoring Procedure

Interrater Reliability

Two raters independently scored 80 of the 242 (33%) scripts. Interrater reliability was high for all scored variables, all absolute agreement intraclass correlations (ICC) ≥ 0.96 based on a two-way random-effects model (see Table 2). Discrepancies between both raters were reviewed and resolved to yield 100% agreement. Given the high interrater agreement, the remaining responses were scored by one rater.

Create *Question Generation Performance*

Following extant scoring procedures (Lim et al., 2024; Wong et al., 2023), participants' *create* question generation test performance was scored by awarding 1 point for each question they produced that fulfilled the *create* level of Bloom's taxonomy. Specifically, *create* questions were those with answers that required creating new knowledge, ideas, or perspectives by compiling information from the study text in a different way, combining elements in a new pattern, or proposing alternative solutions or hypotheses that had not been mentioned in the text (see Table 1). Questions that did not fulfil these criteria were not awarded any points as *create* questions, but were scored as non-*create* questions instead. For instance, questions corresponding to the other levels of Bloom's taxonomy (*remember*, *understand*, *apply*, *analyze*, or *evaluate*) were considered non-*create* questions. Likewise, generic questions that could be produced even without having read and deeply understood the text were considered non-*create* questions (e.g., "What are alternative hypotheses for how the moon was formed?", "What are other reasons why songbirds are vanishing?"). Sample *create* versus non-*create* questions are presented in Table 3.

Table 2 Interrater reliability

Variable	ICC	95% CI
Create question generation performance	.96	[.93, .97]
Comprehension performance	.99	[.990, .996]
Self-other referential terms	1.00	
Elaborations	.97	[.95, .98]
Monitoring statements	.96	[.93, .97]

There was perfect interrater absolute agreement in scoring the number of self-other referential terms. *ICC* intraclass correlation, *CI* confidence interval

Table 3 Sample *create* and non-*create* questions

Question Type	"Origins of the Moon"	"Vanishing Songbirds"
Create	"What new methods can be implemented to ensure that the integrity of the chemical composition of lunar rocks is preserved for better analysis?"	"How would you try to change the natural limitations of breeding season length or egg-laying frequency to increase the tolerance of songbird populations to environmental changes?"
	"Would there be a way for scientists to make the moon a habitable place similar to the earth, since they share some similarities such as the quantities of oxygen isotopes?"	"How can we create a safer environment for songbirds to nest without predators or parasitic birds that can affect their nesting process?"
	"What can we do to find out the actual mass of the impactor which collided with Earth according to the 'giant impact hypothesis', so that we can analyze its relation to the mass of the earth or the moon and also their composition compared to that of the impactor?"	"Are there any natural organisms not currently existing in the ecosystem that can be introduced at a controlled rate such that they could stabilize the songbird population, whilst not becoming an invasive species that is disruptive to the ecosystem?"
Non-*Create*	"Is the moon just a social construct?"	"How should one determine whether songbirds are vanishing or not?"
	"How can we adapt the knowledge about the creation of the moon to understand the creation of other planets?"	"How would the vanishing songbirds affect the ecosystem if their population declines rapidly?"
	"What is the relevance of understanding far objects and astronomical concepts when there are more pressing environmental issues on Earth?"	"To what extent should humans intervene in helping to save songbird populations?"

Comprehension Performance

Each of the six comprehension test questions was scored out of 2 points (i.e., 0 vs. 1 vs. 2), for a total possible score of 12 points. The scoring key is available in the online supplemental materials. For example, for the question "Describe the 'double planet hypothesis' of the origin of the moon. State one scientific fact or research finding supporting that particular theory.", a response would be awarded 1 point for correctly describing the relevant hypothesis, and an additional 1 point for correctly describing a relevant fact or finding that supported the hypothesis.

Process Measures

To examine participants' learning processes during the intervention, their revised study notes, textbook explanations, and verbatim teaching scripts were scored on the number of: (a) self-other referential terms, (b) elaborations, and (c) monitoring statements.

Self-other referential terms were words such as "I", "me", "you", "us", "let's", "our", "ourselves", "we", "your", and "yourself". The number of instances that participants used such terms served as a proxy for their perceived social presence (e.g., Hoogerheide et al., 2016; Jacob et al., 2020, 2021; Lachner et al., 2018; Lim et al., 2021).

Elaborations were statements on the text's content that participants related to their prior knowledge, such as generating analogies and inferences that were not explicitly stated in the text (e.g., Fiorella & Kuhlmann, 2020; Jacob et al., 2020, 2021; Lachner et al., 2018, 2020; Lim et al., 2021). *Monitoring statements* were instances where participants monitored understanding, evaluated correctness, or indicated content that was worth paying attention to based on importance or interest (e.g., Fiorella & Kuhlmann, 2020; Lachner et al., 2020; Roscoe, 2014). Together, elaborations and monitoring statements have been considered elements of reflective knowledge-building that promotes tutors' learning (Roscoe & Chi, 2007). Sample elaborations and monitoring statements are presented in Table 4.

Results

Preliminary Analyses

One-way analyses of variance (ANOVAs) were conducted to ascertain that the three learning groups did not differ at baseline in their pre-studying questionnaire responses, and to explore their post-studying questionnaire responses. Means and standard deviations are presented in Table 5.

Pre-Studying Questionnaire

At baseline, participants reported low perceived prior knowledge of the study topic on overall, with no significant difference across learning groups, $F(2, 239)=0.40$, $p=.67$, $\eta^2 = 0.003$. In addition, the learning groups neither differed in their predicted test performance, $F(2, 239)=0.13$, $p=.88$, $\eta^2 = 0.001$, nor their topic interest ratings, $F(2, 239)=0.13$, $p=.88$, $\eta^2 = 0.001$.

Post-Studying Questionnaire

After the intervention, the learning groups neither differed in their self-reported prior knowledge of the text's content, $F(2, 239)=0.77$, $p=.47$, $\eta^2 = 0.01$, nor how understandable they perceived the text to be, $F(2, 239)=0.30$, $p=.74$, $\eta^2 = 0.003$.

Table 4 Sample elaborations and monitoring statements

Statement Type	Sample Responses
Elaborations	
Analogies	"Imagine playing baseball and your glove catching the ball. Your glove had to be there at the right time and place to catch the ball, just like how the earth had to be right there to catch the moon as it spun by."
	"Nest parasites, similar to nutrient thieves…"
Inferences	"Furthermore, it is known that the moon and earth's compositions are similar, *which suggests that the moon is far more related to the earth than just being an isolated moon in the solar system* [emphasis added]."
	"Long-distance neotropical migrants are believed to decline the most due to tropical deforestation in South America as well—*as their winter homes are destroyed and thus causing harm during the winter periods* [emphasis added]."
Monitoring statements	
Monitoring understanding	"… do you actually know what they are?"
	"… is a little confusing."
Evaluating correctness	"Has it been proven?"
	"… they might or might not be right, just hypotheses."
Directing attention	"It is important to understand that…"
	"This is an interesting fact."

For exploratory purposes, we also examined participants' ratings on the eight phenomenological items. The learning groups' stress ratings differed, $F(2, 239) = 4.79$, $p = .01$, $\eta^2 = 0.04$, whereby the teaching group reported higher stress when learning about the subject matter than the notetaking and explaining groups, both $ps = .01$, $d = 0.43$ and 0.40, respectively; the latter two groups did not differ, $p = .88$, $d = 0.03$. The learning groups also differed in the amount of mental effort that they reported investing when learning the subject matter, $F(2, 239) = 3.85$, $p = .02$, $\eta^2 = 0.03$. Specifically, the teaching group reported investing more mental effort than the notetaking group, $p = .01$, $d = 0.43$, but not the explaining group, $p = .13$, $d = 0.25$; the latter two groups did not differ, $p = .22$, $d = 0.19$. There were no other significant differences across learning conditions on the remaining phenomenological items, all $ps > .05$.

Main Analyses

Total Questions Generated

Because participants were instructed to generate as many *create* questions as possible at test, we ascertained that the total number of questions that they attempted to generate (including questions that actually fulfilled the *create* level versus non-*create* questions that did not) did not differ across learning conditions (e.g., Lim et al., 2024).

Table 5 Means and standard deviations for pre-studying, post-studying, and post-experiment questionnaires

Variables	Notetaking		Explaining		Teaching	
	M	SD	M	SD	M	SD
Pre-studying questionnaire						
Perceived prior knowledge of topic	1.32	0.65	1.29	0.66	1.23	0.55
Predicted test performance	2.15	0.96	2.21	0.82	2.21	0.89
Topic interest	2.99	0.75	2.97	0.92	3.04	0.86
Post-studying questionnaire						
Post-studying JOL	50.99	20.59	49.12	21.89	47.65	22.82
Perceived prior knowledge of text	1.37	0.86	1.48	1.08	1.57	1.10
Text understandability	5.07	1.39	5.14	1.15	4.98	1.45
Difficult	3.63	1.36	3.48	1.55	3.54	1.55
Enjoyment	4.63	1.50	4.65	1.54	4.63	1.44
Future	4.23	1.41	4.28	1.49	4.17	1.50
Understand	3.89	1.38	4.11	1.46	4.12	1.44
Interest	3.85	1.53	3.71	1.65	4.00	1.67
Useful	3.30	1.56	3.03	1.53	3.46	1.83
Stress	3.28	1.63	3.33	1.64	4.01	1.77
Effort	4.62	1.21	4.85	1.17	5.14	1.19
Post-experiment questionnaire						
Method effectiveness	3.85	1.46	4.43	1.22	4.07	1.44

N = 242, *JOL* judgment of learning

Indeed, a 3 (learning method) × 2 (study topic) between-subjects ANOVA indicated that there was no significant difference across the notetaking (M=5.51, SD=1.95), explaining (M=5.83, SD=2.51), and teaching (M=5.47, SD=2.75) groups in their total questions generated, $F(2, 236)=0.52$, p=.60, $\eta_p^2 = 0.004$. Thus, any differences in students' generation of, specifically, *create* questions cannot be attributed to more (or fewer) attempts in any particular learning condition.

Overall, participants generated more questions for the "vanishing songbirds" (M=6.08, SD=2.58) than "origins of the moon" (M=5.13, SD=2.17) topic, $F(1, 236)=9.45$, p=.002, $\eta_p^2 = 0.04$. Nevertheless, the learning method × study topic interaction was non-significant, $F(2, 236)=1.84$, p=.16, $\eta_p^2 = 0.02$.

Create *Question Generation Performance*

Focusing on the number of *create* questions that participants generated at test, a 3 (learning method) × 2 (study topic) between-subjects ANOVA revealed a significant difference across learning groups, $F(2, 236)=18.23$, p<.001, $\eta_p^2 = 0.13$. As predicted, the explaining (M=1.54, SD=1.67) and teaching (M=2.16, SD=1.93) groups outperformed the notetaking (M=0.75, SD=1.06) group, p=.001 and p<.001, d=0.57 and 0.91, respectively. Importantly, the teaching group generated more *create* questions than the explaining group, p=.008, d=0.34. Thus, writing textbook explanations or teaching scripts enhanced participants' *create* question generation performance more than writing study notes for their own learning, with an additional benefit from teaching than explaining (Fig. 2A).

Overall, participants generated more *create* questions for the "vanishing song-birds" ($M = 2.06$, $SD = 1.90$) than "origins of the moon" ($M = 0.92$, $SD = 1.22$) topic, $F(1, 236) = 35.35, p < .001, \eta_p^2 = 0.13$. Nevertheless, the learning method × study topic interaction was non-significant, $F(2, 236) = 1.21, p = .30, \eta_p^2 = 0.01$, indicating that the advantages of explaining and teaching generalized across both topics.

Comprehension Performance

A 3 (learning method) × 2 (study topic) between-subjects ANOVA indicated that the learning groups significantly differed in their comprehension test performance,

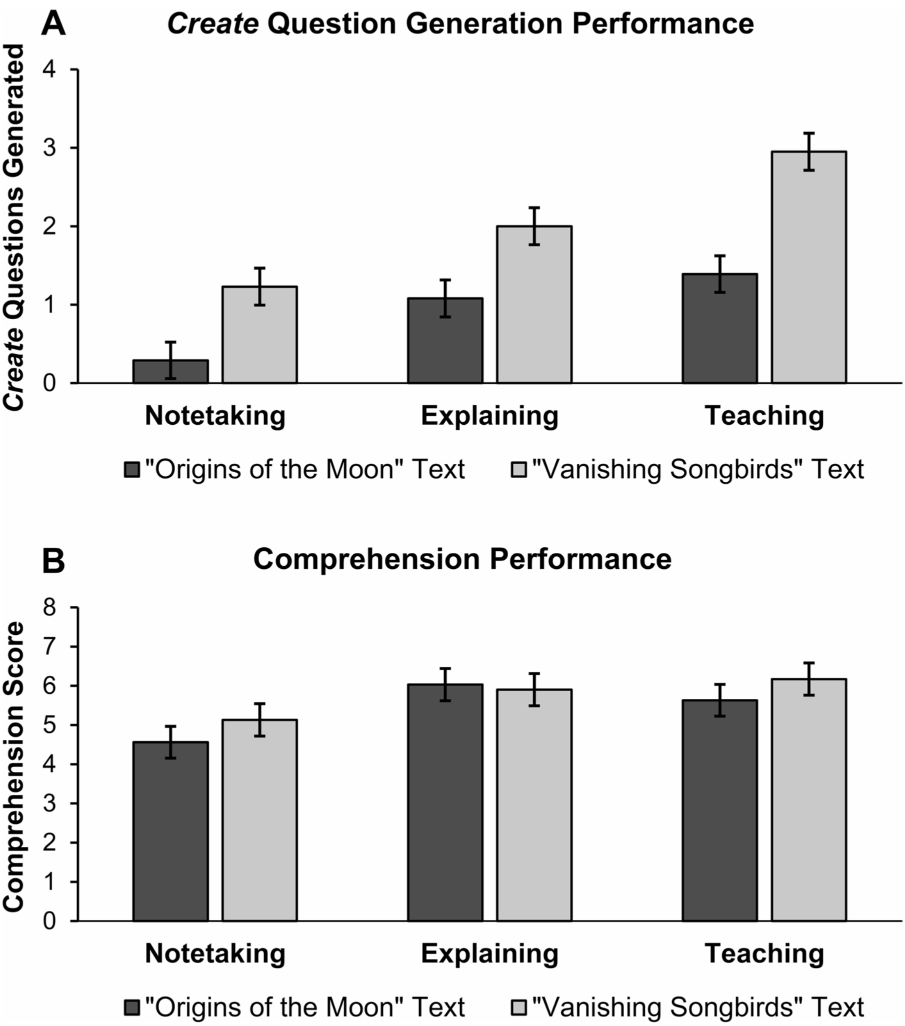

Fig. 2 *Create* question generation and comprehension test performance across learning conditions and study topics

$F(2, 236) = 4.75$, $p = .01$, $\eta_p^2 = 0.04$, albeit with a different pattern from their *create* question generation performance. Both the explaining ($M = 5.96$, $SD = 2.49$) and teaching ($M = 5.90$, $SD = 2.82$) groups displayed better comprehension performance than the notetaking ($M = 4.84$, $SD = 2.45$) group, $p = .007$ and .01, $d = 0.45$ and 0.40, respectively. Interestingly, though, the explaining and teaching groups did not significantly differ in their comprehension performance, $p = .89$, $d = -0.02$. Thus, generating materials for others' learning—whether by writing textbook explanations or teaching scripts—enhanced comprehension of the material, relative to generating study notes for one's own learning (Fig. 2B).

Overall, participants' comprehension test performance did not significantly differ across the "origins of the moon" ($M = 5.40$, $SD = 3.05$) and "vanishing songbirds" ($M = 5.73$, $SD = 2.13$) topics, $F(1, 236) = 0.96$, $p = .33$, $\eta_p^2 = 0.004$. Neither was there a learning method × study topic interaction, $F(2, 236) = 0.46$, $p = .64$, $\eta_p^2 = 0.004$.

Given that the explaining and teaching groups did not differ in their superior comprehension performance over the notetaking group, it is unlikely that the teaching advantage for *create* question generation was simply driven by better comprehension of the text content. To further ascertain this, participants' comprehension performance was added as a covariate in a 3 (learning method) × 2 (study topic) analysis of covariance (ANCOVA) with their *create* question generation performance as the dependent variable. To first check the homogeneity of slopes assumption, interactions between the covariate and independent variables were entered into the model, alongside the main effects. None of the interactions were significant, all ps>.05, indicating that the homogeneity of slopes assumption had been met. The interaction terms were then removed from the model, which was re-estimated.

The ANCOVA indicated that the teaching advantage for *create* question generation persisted even after controlling for comprehension performance, $F(2, 235) = 15.74$, $p < .001$, $\eta_p^2 = .12$, whereby the teaching group ($M_{adjusted} = 2.14$) generated significantly more *create* questions than the notetaking ($M_{adjusted} = 0.82$) and explaining ($M_{adjusted} = 1.50$) groups, $p < .001$ and $p = .007$, respectively. In addition, the explaining group outperformed the notetaking group, $p = .004$. This suggests that the teaching and explaining groups' better *create* question generation performance was not merely due to their better comprehension of the text per se.

Metacognitive Judgments

In contrast to their actual test performance, participants' metacognitive judgments revealed that they were largely unaware of the benefits of teaching. Means and standard deviations are presented in Table 5. After the intervention, the three learning groups did not significantly differ in their post-studying JOL ratings, $F(2, 239) = 0.48$, $p = .62$, $\eta^2 = 0.004$. After experiencing the methods' effects on their test performance, though, participants' ratings of their method's effectiveness differed across learning conditions, $F(2, 239) = 3.54$, $p = .03$, $\eta^2 = 0.03$. Specifically, the explaining group correctly judged their learning method as more effective than the notetaking group, $p = .009$, $d = 0.43$. However, the teaching group's method effectiveness ratings did not significantly differ from those of the notetaking and explaining groups, $p = .31$ and .11, $d = 0.15$ and -0.27, respectively.

Process Measures

To examine the processes that participants engaged in during the intervention, the number of self-other referential terms, elaborations, and monitoring statements in their revised study notes, textbook explanations, and teaching scripts were analyzed, as was their self-reported sense of power after the intervention. Means and standard deviations are presented in Table 6.

Self-Other Referential Terms

As a proxy for perceived social presence, the number of self-other referential terms in participants' studying responses was analyzed, revealing a significant difference across learning conditions, $F(2, 239) = 36.72$, $p < .001$, $\eta^2 = 0.24$. The teaching group used more self-other referential terms than the notetaking and explaining groups, both *p*s<.001, $d = 1.11$ and 0.90, respectively. In contrast, the notetaking and explaining groups did not significantly differ, $p = .19$, $d = 0.36$. Thus, writing a verbatim teaching script triggered greater perceived social presence than writing study notes or textbook explanations.

Elaborations

The learning groups also differed in the number of elaborations in their studying responses, $F(2, 239) = 10.39$, $p < .001$, $\eta^2 = 0.08$, albeit with a different pattern. Both the explaining and teaching groups generated more elaborations than the notetaking group, $p = .004$ and $p < .001$, $d = 0.51$ and 0.73, respectively. However, the explaining and teaching groups did not significantly differ in their number of elaborations, $p = .11$, $d = 0.22$. Hence, writing a textbook explanation or verbatim teaching script induced more generative processing than writing study notes.

Monitoring Statements

The number of metacognitive monitoring statements in participants' studying responses differed across learning conditions, $F(2, 239) = 15.32$, $p < .001$, $\eta^2 = 0.11$. As with their use of self-other referential terms, the teaching group generated more monitoring statements than the notetaking and explaining groups, both *p*s<.001, $d = 0.76$ and 0.57, respectively. In contrast, the notetaking and explaining groups did not significantly differ, $p = .22$, $d = 0.25$. Thus, writing verbatim teaching scripts

Table 6 Means and standard deviations for self-other referential terms, elaborations, monitoring statements, and sense of power

Variables	Notetaking		Explaining		Teaching	
	M	SD	M	SD	M	SD
Self-other referential terms (perceived social presence)	0.32	0.97	0.73	1.28	2.77	2.96
Elaborations	0.54	0.87	1.12	1.37	1.44	1.50
Monitoring statements	0.30	0.77	0.50	0.81	1.16	1.41
Sense of power	3.75	1.20	4.18	1.12	4.59	1.14

induced greater metacognitive monitoring than writing study notes or textbook explanations.

Sense of Power

Finally, participants' mean sense of power ratings were analyzed, revealing a significant difference across learning conditions, $F(2, 239) = 10.71$, $p < .001$, $\eta^2 = 0.08$. Both the explaining and teaching groups reported experiencing more power during the intervention than the notetaking group, $p = .018$ and $p < .001$, $d = 0.37$ and 0.72, respectively. Crucially, the teaching group reported a greater sense of power than the explaining group, $p = .027$, $d = 0.36$. Thus, parallel to participants' *create* question generation performance, writing a textbook explanation or teaching script activated greater feelings of power than writing study notes, with a further boost for teachers than explainers.

Process Measures and Test Performance

Correlations

Correlational analyses were conducted to examine associations between participants' learning processes during the intervention and their subsequent test performance. The correlation matrix is displayed in Table 7. As shown in the matrix, participants' *create* question generation performance significantly and positively correlated with their sense of power and use of self-other referential terms, elaborations, and monitoring statements during the intervention. However, only participants' sense of power and use of elaborations significantly and positively correlated with their comprehension test performance.

Mediation Analyses: Power Hypothesis

All in, the preceding results showed that teaching produced better *create* question generation performance than explaining, which in turn outperformed notetaking. Furthermore, experiencing a heightened sense of power during the intervention was significantly associated with generating more *create* questions at test. Parallel to participants' *create* question generation performance, teaching induced a greater sense of power than explaining, which in turn induced more power than notetaking. In

Table 7 Correlations for test performance and process measures

Variables	1	2	3	4	5	6
1. *Create* question generation performance	—					
2. Comprehension performance	.21**	—				
3. Self-other referential terms	.26***	−.03	—			
4. Elaborations	.14*	.18**	.28***	—		
5. Monitoring statements	.25***	.10	.52***	.28***	—	
6. Sense of power	.23***	.18**	.28***	.19**	.19**	—

*$p<.05$. **$p<.01$. ***$p<.001$

contrast, this stepwise pattern did not occur for all other process measures during the intervention: self-other referential terms, elaborations, and monitoring statements. Accordingly, regression analyses were run to test the power hypothesis of learning-by-teaching—did a greater sense of power mediate the benefit of teaching for *create* question generation?

Following Hayes and Preacher's (2014) guidelines, a percentile bootstrap estimation approach with 10,000 samples was used in Model 4 of Hayes's (2013) PROCESS macro with learning method as the multicategorical predictor, sense of power as the mediator, and *create* question generation performance as the outcome. Using Helmert coding, the respective contrast codes assigned to the notetaking, explaining, and teaching conditions were: −2/3, 1/3, and 1/3 for the first contrast (X_1), and 0, −1/2, and 1/2 for the second contrast (X_2). Thus, X_1 tested whether sense of power mediated the effect of generating materials for others' learning (explaining and teaching conditions) on *create* question generation performance, compared to generating materials for one's own learning (notetaking condition). X_2 tested whether sense of power mediated the effect of teaching on *create* question generation performance, compared to explaining.

Figure 3 depicts the mediation model. In mediation analyses with a multicategorical predictor, evidence that at least one relative indirect effect differs from zero supports the conclusion that a mediator variable mediates the effect of the predictor on the outcome (Hayes & Preacher, 2014). For X_1 that compared notetaking against the explaining and teaching groups, the relative indirect effect was significant, 0.13, 95% CI [0.02, 0.29]. Likewise, for X_2 that compared explaining versus teaching, the relative indirect effect was significant, 0.08, 95% CI [0.002, 0.21]. Hence, consistent with the power hypothesis, a greater sense of power mediated the benefit of teaching over explaining, as well as the benefit of explaining and teaching over notetaking, for *create* question generation performance.[1]

Discussion

Teaching others is a powerful way to improve one's own learning, particularly for higher order outcomes. The present study found that generating materials for others' learning—writing a verbatim teaching script or textbook explanation—of a scientific text enabled students to ask more research questions that create new knowledge about the material, relative to writing study notes for their own learning. Crucially, teaching produced an additional boost over explaining for students' *create* question generation performance. At the same time, writing teaching scripts or textbook explanations enhanced students' comprehension of the material more than writing study notes,

[1] To further ascertain that the other three process measures did not account for the effect of learning method on *create* question generation performance, we separately conducted similar mediation analyses with self-other referential terms, elaborations, and monitoring statements as the mediator, respectively. Indeed, the analyses affirmed that none of these three process measures mediated the teaching advantage for *create* question generation performance—all relative indirect effects were non-significant with zero included in their respective 95% CIs. Thus, we found no evidence for the social presence hypothesis and generative hypothesis in accounting for the teaching advantage in *create* question generation.

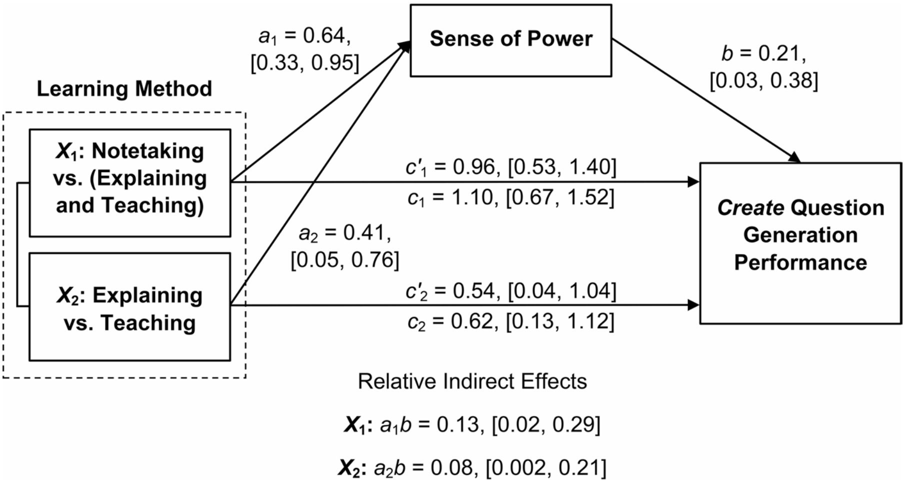

Fig. 3 Results of mediation analyses. Note: Mediation model with learning method as the multicategorical predictor, sense of power as the mediator, and *create* question generation performance as the outcome. Using Helmert coding, the respective contrast codes assigned to the notetaking, explaining, and teaching conditions were: -2/3, 1/3, and 1/3 for the first contrast (X_1), and 0, -1/2, and 1/2 for the second contrast (X_2). Unstandardized regression coefficients are presented for each path. The c' path coefficients represent the relative direct effects, whereas the c path coefficients represent the relative total effects. Values in brackets represent the 95% confidence intervals for the regression coefficients using a percentile bootstrap estimation approach with 10,000 samples.

with no significant difference between the former two methods. Thus, acquiring better basic comprehension of to-be-learned material does not necessarily translate to gains for more complex, higher order learning outcomes (e.g., Agarwal, 2019; Wong et al., 2023). Notably, the advantage of teaching for *create* question generation held even after controlling for students' comprehension test performance, indicating that it was not merely driven by an improved understanding of the text.

Yet, students' metacognitive judgments revealed that they were largely unaware of the benefits of teaching. Indeed, after the intervention, students inaccurately perceived teaching to be just as effective as notetaking and explaining. This metacognitive illusion persisted even after students had experienced the learning methods' effects on their test performance, although they correctly judged explaining as more effective than notetaking.

Overall, these findings add to extant research on the benefits of generating materials for others' learning over egocentric learning activities (e.g., Ribosa & Duran, 2022; Wong et al., 2023), while going beyond it to demonstrate the unique advantage of teaching over explaining for complex learning outcomes such as research question generation. Thus, the intentionality of teaching—taking on the role of a teacher—carries weight in determining the extent that students benefit from generating learning materials for others (Kobayashi, 2023).

Why does teaching others enhance one's own learning? This study advanced and found evidence for a novel power hypothesis of learning-by-teaching. When students took on the role of a teacher in writing teaching scripts, they experienced more power in influencing others than writing textbook explanations, which in turn induced more power than writing study notes. This heightened sense of power then mediated the advantage of teaching over explaining, as well as that of teaching and explaining over notetaking, for students' *create* question generation performance.

Theoretical Implications

Power in Learning-by-Teaching

As the first demonstration of the power hypothesis of learning-by-teaching, the present findings contribute several theoretical insights. Whereas extant accounts of learning-by-teaching have focused on how teaching triggers increased feelings of social presence in perceiving one's audience as "real" and "present" (e.g., Hoogerheide et al., 2016; Jacob et al., 2020), we show that teaching potentially induces an increased sense of power in influencing others. This finding aligns with the notion that interpersonal behaviors shape the psychological experience of power (Tost, 2015), and establishes teaching as a possible account of this experience. Consistent with extant work on advice-giving (Schaerer et al., 2018), our data reveal that the mere act of taking on the role of a teacher and teaching others—even without interacting with one's audience—is sufficient to instill in students a sense of power.

In turn, an increased sense of power potentially benefits the "teacher's" own deep learning, in this case assessed as their ability to generate more high-quality research questions that create new knowledge about the learned material. Stepping into a position of power may induce in powerholders a sense of responsibility for the outcomes of those under their charge (Anderson & Galinsky, 2006; Overbeck & Park, 2001), which could lead powerholders to be more responsive to others' needs (Tost, 2015; Tost & Johnson, 2019). Moreover, a sense of power promotes an approach orientation, leading powerholders to enact agentic behaviors (Galinsky et al., 2003; Tost, 2015). In learning-by-teaching contexts, feeling powerful in the role of a teacher could thus motivate students to improve their own deep learning and master the material to teach their audience well.

Whereas the power hypothesis has been proposed and unveiled for the first time here, this account is compatible with other recent findings in learning-by-teaching research. For instance, it can explain why learning-by-teaching effects vary depending on the social identity or status of one's audience. In a recent study, Pi et al. (2025) found that students learned better after orally explaining an educational video to a less knowledgeable peer than a knowledgeable teacher, especially when the students had low prior knowledge. The authors interpreted these findings in favor of the generative hypothesis and social presence hypothesis, suggesting that the students may have adapted their explanations by generating more elaborations for a less knowledgeable peer than a knowledgeable teacher. But as the authors also acknowledged, they did not analyze the quality of students' explanations (e.g., number of elaborations), such that no evidence could be marshalled for these hypotheses. Concurrently, we note that this interpretation cannot fully account for why students with high prior knowledge would not also have generated more elaborations when explaining to a peer than a teacher. Rather, these findings can be interpreted through the lens of the power hypothesis. Being in an objective position of power does not always implicate subjectively feeling powerful (Smith & Hofmann, 2016), especially when the powerholder has low self-perceived legitimacy and/or lacks confidence in their abilities to perform up to expectations (Tost, 2015). Hence, particularly when Pi et al.'s (2025) students had low prior knowledge, they may have experienced a greater sense of power and thus learned better from explaining to a less knowledgeable peer, as opposed to a knowledgeable teacher whom the students may further have (implicitly) perceived as assessing, rather than learning from, their explanation.

Relatedly, Wang et al. (2023) found that students performed better on a transfer test after they had taught an imaginary audience while being video-recorded, relative to teaching another student face-to-face or teaching a group of seven students face-to-face. Of note, students who taught an imaginary audience reported lower social presence, experienced lower physiological arousal as measured by pulse rate, and engaged in more generative processing than those who taught a physically present audience. Moreover, teaching an imaginary audience produced lower self-reported state anxiety and cognitive load than teaching a group of students face-to-face. Thus, contrary to the social presence hypothesis of learning-by-teaching, *minimizing* the audience's social presence in fact maximized learning outcomes. Wang et al. (2023) interpreted their findings as boundary conditions of the social presence hypothesis, whereby too much social presence is distracting to the point of daunting. Alternatively, we note that these findings are consistent with how anxiety blocks the psychological effects of power on agency and approach (Maner et al., 2012). Anxiety has been linked to behavioral inhibition and a focus on threat (Keltner et al., 2003; MacLeod & Mathews, 1988), such that anxious people may prioritize avoiding errors or potential risks, undermining their tendency to act in agentic and approach-oriented ways (Maner et al., 2012). Accordingly, the potential benefits of power when students take on the role of a teacher could be moderated by their anxiety levels.

Of course, power is not necessarily the only mechanism for learning-by-teaching effects; other processes could play contributing roles too. For instance, in line with the generative hypothesis, some studies have found that greater generative processing in producing more elaborations mediates the benefits of learning-by-teaching for transfer (Lachner et al., 2018) and holistic argumentation quality (Wong, 2025). In the present study, though, such processing did not suffice to account for the teaching advantage in research question generation, as will be discussed next.

Teaching Versus Explaining

The terms "teaching" and "explaining" have often been used interchangeably in extant learning-by-teaching (or -explaining) literature. Indeed, both similarly involve generating materials for others' learning. However, this study demonstrates that teaching versus explaining are distinct activities with differential impacts on learning processes and outcomes. For instance, students displayed higher social presence when writing teaching scripts than textbook explanations or study notes, as measured by their use of more self-other referential terms. Conversely, explaining did not increase social presence more than notetaking, even though students' textbook explanations were directed at their peers as the intended audience whereas their study notes were directed at themselves. A similar pattern occurred in students' use of monitoring statements, whereby teaching increased metacognitive monitoring more than explaining and notetaking, whereas the latter two methods did not differ. Thus, neither social presence nor metacognitive monitoring could fully account for why explaining subsequently produced better *create* question generation performance than notetaking.

In addition, teaching and explaining were both just as effective in triggering more elaborations than notetaking. Thus, it is unlikely that generative processing alone drove the learning advantage of teaching over explaining in *create* question generation. Moreover, the teaching benefit unlikely hinged on retrieval processes since all students engaged in open-book studying during the intervention, such that there was little basis for them to engage in retrieval practice. Taken together, extant accounts of learning-by-teaching—the social presence hypothesis, generative hypothesis, and retrieval hypothesis—are limited in fully explaining the present set of findings, although this could in part be because the operationalizations of social presence and generative and monitoring processes in this study are necessarily limited, given that the measures relied on particular proxies such as the frequency of self-other referential terms or elaborations and monitoring statements, such that the present results cannot definitively rule out these mechanisms across other tasks and measures.

Educational Implications

The traditional student–teacher power relationship has been characterized as one marked by a student's deference to a teacher's authoritative knowledge and power to affirm the student and build self-esteem (Symonds, 2021). As this study demonstrates, however, when students are empowered to *be* the teacher, they learn better. Here, it is important to note that teaching does not simply involve explaining to-be-learned material. Whereas all teachers explain, not all who explain are teachers. Although having students write a textbook explanation for their peers can benefit their learning, students gain further from specifically taking on the role of a teacher to write a teaching script, just as how they would orate a lecture to their peers. Based on recent evidence that learning-by-teaching benefits are further amplified when students teach incorrectly with deliberate errors that they have generated for their intended audience to spot (i.e., *learning by misteaching*; Wong, 2025; see also Wong, 2023; Wong & Lim, 2022a, 2022b; Yap & Wong, 2024), students could also be guided to strategically incorporate errors in their teaching for optimized learning (Wong & Lim, 2019). Moreover, writing teaching scripts has been found to boost learning outcomes not only in lab settings as in this and other studies (Lim et al., 2021; Wong et al., 2023), but also in real-world classrooms (Lim et al., 2024). Hence, the intentionality of teaching should not be trivialized, even though student-generated textbook explanations and teaching scripts can both serve as learning resources if shared with their peers (Kobayashi, 2023).

However, students are often unaware of the benefits of teaching, as evident from their inaccurate metacognitive judgments in this and other studies (e.g., Lim et al., 2021; Wong, 2025). Indeed, people are often unaware of how possessing power alters their thinking, emotions, and behavior (Smith & Galinsky, 2010). Furthermore, recent evidence suggests that learners who feel less knowledgeable about the material are more reluctant to engage in explaining it to others (Atir & Risen, 2025). Yet, taking on the role of a teacher to teach could paradoxically imbue such learners with a greater sense of power in influencing others, in turn potentially benefiting their own learning. Thus, increasing students' beliefs about the effectiveness of learning-by-teaching could be vital for encouraging them to choose to use and profit from this technique (Atir & Risen, 2025).

Limitations and Future Directions

Building on the present initial evidence, much more room remains to further test, develop, and refine the novel power hypothesis. Arising from the lack of whole measures that assess power in learning-by-teaching, this study developed a sense-of-power measure by extracting items from extant established measures that specifically relate to power as a sense of one's capacity to influence others (Anderson et al., 2012; Cassidy & Lynn, 1989; Schaerer et al., 2018), and adapting these items for the learning-by-teaching context. As a promising starting point, the present sense-of-power measure demonstrated high internal consistency (Cronbach's $\alpha = .83$). Future work ought to examine the measure's construct validity (e.g., convergent and discriminant validity) and generalizability across task types, learner populations, and teaching situations. It is our hope that the present measure will serve as the foundation for spurring new interest within the community toward developing and validating scales of power in a learning-by-teaching context, eventually enabling the optimal measurement of this construct.

A worthwhile endeavor could aim to further differentiate the present notion of power from its possible correlates, in turn concretizing a heightened sense of power as an independent account of the observed benefits in learning-by-teaching. This study's findings go some way in ruling out some possible correlates. Participants' ratings of their subjective learning experience revealed that the three learning groups did not differ in their enjoyment or interest in learning more about the subject matter during the intervention, suggesting that perceived learning enjoyment or interest did not produce the observed benefits of teaching or explaining (cf. Hoogerheide et al., 2019b). Moreover, although the teaching group reported experiencing higher stress and investing more mental effort when learning about the subject matter, the notetaking and explaining groups did not differ on these fronts. Thus, perceived stress and mental effort cannot parsimoniously account for the present learning benefits (see Hoogerheide et al., 2019b for similar findings on the non-contributions of perceived effort investment), given that explainers still outperformed notetakers on both the *create* question generation and comprehension tests. Beyond these variables, though, an increased sense of power may be motivating, and could promote a mental state focused on goal-directed actions (Guinote, 2007a) and judicious self-regulation to achieve those goals (DeWall et al., 2011; Guinote, 2007b). Future examinations of power in learning-by-teaching could consider such potential correlates as motivation and cognitive control.

Relative to explaining, teaching bolstered higher order *create* question generation performance, although teaching and explaining did not differ for basic comprehension performance. That the present power hypothesis accounts particularly for higher order, but not necessarily basic, learning outcomes is an open avenue for future inquiry. Furthermore, even for higher order learning outcomes, it would be premature to assume that power necessarily exerts unconditional main effects in learning-by-teaching. Two famous sayings about the dark versus bright side of power aptly illustrate the tension between the notions that "power tends to corrupt" versus "with great power comes great responsibility". A wealth of social power research has documented seemingly contradictory effects, in part because how people respond to power depends on not only whether they experience a high sense of power, but also how they construe—cognitively appraise and assign meaning to—power and its implications (Scholl et al., 2022; see also Kipnis, 1972). Depending on the situation or cultural context, powerholders can construe power as an *opportunity* to freely "make things happen" or as a *responsibility* to "take care of things" that only they can manage, as with the notion of "noblesse oblige" (Sassenberg et al., 2012; Scholl, 2020; Scholl et al., 2022). In learning-by-teaching, teachers may construe their power as an opportunity to freely decide what and how their students should learn versus as responsibility in ensuring that their students learn well. In turn, these different construals of power alter powerholders' behavior. Whereas construing power as opportunity may promote a narrow focus on achieving the goal at hand even through selfish behaviors, construing power as responsibility may promote a broader and more deliberative approach, such as taking others' advice rather than disregarding it (De Wit et al., 2017). Thus, how people construe power in the role of a teacher could potentially influence how they prepare to teach, their actual teaching style, and their subsequent learning.

Besides varying construals of power, personality traits such as openness to experience could moderate the impact of sense of power on creativity. For instance, Magni et al. (2024) found that one's sense of power has countervailing effects in enhancing creativity through increased risk-taking while hindering creativity through reduced perspective-taking. However, for individuals with high openness, the positive pathway through risk-taking is amplified over the negative pathway through perspective-taking, such that a sense of power exerts a more positive effect on these individuals' creativity on overall. When and for whom does the power hypothesis of learning-by-teaching apply? Exploring these nuances presents exciting avenues for future work in contributing to theory development.

Concurrently, it would be valuable to examine how students can be prompted or guided to teach more effectively for potentially greater learning gains. Here, we observed that participants' study responses contained relatively few elaborations and monitoring statements, likely because participants taught spontaneously without any explicit training (e.g., Fiorella & Kuhlmann, 2020; Wong, 2025). Yet, lower teaching quality could constrain the teacher's learning (Roscoe, 2014; Roscoe & Chi, 2007). Thus, future work that applies insights from effective teacher professional development to learning-by-teaching could potentially support sustained improvements in students' teaching and learning. For instance, effective training could provide students with insight on how teaching and learning work, build their motivation to adopt effective practices, equip them with the necessary techniques to enact their insights, and incorporate practice of such techniques to firmly root them in students' repertoire (Sims et al., 2025). As the advance of generative artificial intelligence (AI) rapidly (re)shapes education, it would also be valuable to examine how students can effectively collaborate with AI during their teaching to reap learning gains for their independent creativity (Wong & Qiu, 2026).

Finally, a summary—delineation—of the present paradigm's features deserves mention. In this study, the teaching was non-interactive and conducted in written (rather than oral) form, learners studied relatively short expository texts in an open-book fashion, the learning outcome of main interest was *create*-level research question generation, and the sample comprised university students in a large public English-speaking university. We encourage empirical investigations in systematically assessing the present power hypothesis's applicability in other forms of learning-by-teaching, and welcome collaborations.

Conclusion

When students teach, they learn—powerfully. This study further advanced a novel power hypothesis of learning-by-teaching. More than just writing textbook explanations for their peers or study notes for their own learning, our results suggest that when students step into the role of a teacher and write teaching scripts as how they would orate a lecture for their peers, they may gain a heightened sense of power in influencing others, in turn potentially benefiting their research question generation performance. Indeed, teaching is powerful.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10648-025-10097-1.

Acknowledgements The authors thank Brittney Zi Qing Loke and Megan Ruyun Tham for their assistance with data collection and/or scoring.

Author contributions Sarah Shi Hui Wong served as lead for formal analysis, funding acquisition, and writing–original draft, contributed equally to conceptualization, methodology, and resources, and served in a supporting role for writing–review and editing. Stephen Lim served as lead for writing–review and editing, and contributed equally to conceptualization, methodology, and resources.

Funding This research was supported by a Singapore Management University grant and a Yale-NUS College Research grant (H-607-02-000047) awarded to Sarah Shi Hui Wong.

Materials Availability The materials for this study are available in the online supplementary materials.

Declarations

Ethics Approval This research was conducted with the appropriate ethics approval from the National University of Singapore's institutional review board.

Consent To Participate Informed consent was obtained from all participants included in the study.

Competing Interests The authors have no relevant financial or non-financial interests to disclose.

References

Agarwal, P. K. (2019). Retrieval practice & Bloom's taxonomy: Do students need fact knowledge before higher order learning? *Journal of Educational Psychology, 111*(2), 189–209. https://doi.org/10.1037/edu0000282

Allen, V. L., & Feldman, R. S. (1973). Learning through tutoring: Low-achieving children as tutors. *The Journal of Experimental Education*, *42*(1), 1–5. https://doi.org/10.1080/00220973.1973.11011433

Anderson, C., & Galinsky, A. D. (2006). Power, optimism, and risk-taking. *European Journal of Social Psychology*, *36*(4), 511–536. https://doi.org/10.1002/ejsp.324

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives* (abridged ed.). Addison Wesley Longman, Inc.

Anderson, C., John, O. P., & Keltner, D. (2012). The personal sense of power. *Journal of Personality*, *80*(2), 313–344. https://doi.org/10.1111/j.1467-6494.2011.00734.x

Atir, S., & Risen, J. L. (2025). The paradox of explaining: When feeling unknowledgeable prevents learners from engaging in effective learning strategies. *Journal of Experimental Psychology: General*, *154*(1), 228–248. https://doi.org/10.1037/xge0001679

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

Bargh, J. A., & Schul, Y. (1980). On the cognitive benefits of teaching. *Journal of Educational Psychology*, *72*(5), 593–604. https://doi.org/10.1037/0022-0663.72.5.593

Benware, C. A., & Deci, E. L. (1984). Quality of learning with an active versus passive motivational set. *American Educational Research Journal*, *21*(4), 755–765. https://doi.org/10.3102/00028312021004755

Biswas, G., Leelawong, K., Schwartz, D., Vye, N., & The Teachable Agents Group at Vanderbilt. (2005). Learning by teaching: A new agent paradigm for educational software. *Applied Artificial Intelligence, 19*(3–4), 363–392. https://doi.org/10.1080/08839510590910200

Biswas, G., Segedy, J. R., & Bunchongchit, K. (2016). From design to implementation to practice a learning by teaching system: Betty's brain. *International Journal of Artificial Intelligence in Education*, *26*(1), 350–364. https://doi.org/10.1007/s40593-015-0057-9

Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals*. McKay.

Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(6), 1563–1569. https://doi.org/10.1037/a0017021

Cassidy, T., & Lynn, R. (1989). A multifactorial approach to achievement motivation: The development of a comprehensive measure. *Journal of Occupational Psychology*, *62*(4), 301–312. https://doi.org/10.1111/j.2044-8325.1989.tb00001.x

Cheng, M., Wang, F., & Mayer, R. E. (2023). Benefits of asking students to make an instructional video of a multimedia lesson: Clarifying the learning-by-teaching hypothesis. *Journal of Computer Assisted Learning, 39*(5), 1636–1651. https://doi.org/10.1111/jcal.12823

Chin, C., & Brown, D. E. (2002). Student-generated questions: A meaningful aspect of learning in science. *International Journal of Science Education*, *24*(5), 521–549. https://doi.org/10.1080/09500690110095249

Chin, D. B., Dohmen, I. M., Cheng, B. H., Oppezzo, M. A., Chase, C. C., & Schwartz, D. L. (2010). Preparing students for future learning with teachable agents. *Educational Technology Research and Development*, *58*(6), 649–669. https://doi.org/10.1007/s11423-010-9154-5

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B.Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127–149). American Psychological Association. https://doi.org/10.1037/10096-006

Clinton, V., & van den Broek, P. (2012). Interest, inferences, and learning from texts. *Learning and Individual Differences*, *22*(6), 650–663. https://doi.org/10.1016/j.lindif.2012.07.004

De Wit, F. R. C., Scheepers, D., Ellemers, N., Sassenberg, K., & Scholl, A. (2017). Whether power holders construe their power as responsibility or opportunity influences their tendency to take advice from others. *Journal of Organizational Behavior*, *38*(7), 923–949. https://doi.org/10.1002/job.2171

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

DeWall, C. N., Baumeister, R. F., Mead, N. L., & Vohs, K. D. (2011). How leaders self-regulate their task performance: Evidence that power promotes diligence, depletion, and disdain. *Journal of Personality and Social Psychology*, *100*(1), 47–65. https://doi.org/10.1037/a0020932

Dillon, J. T. (1984). The classification of research questions. *Review of Educational Research*, *54*(3), 327–361. https://doi.org/10.3102/00346543054003327

Duran, D., & Topping, K. (2017). *Learning by teaching: Evidence-based strategies to enhance learning in the classroom*. Routledge. https://doi.org/10.4324/9781315649047

Eskreis-Winkler, L., Fishbach, A., & Duckworth, A. L. (2018). Dear Abby: Should I give advice or receive it? *Psychological Science*, *29*(11), 1797–1806. https://doi.org/10.1177/0956797618795472

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fiorella, L. (2023). Making sense of generative learning. *Educational Psychology Review, 35*(2), Article 50. https://doi.org/10.1007/s10648-023-09769-7

Fiorella, L., & Kuhlmann, S. (2020). Creating drawings enhances learning by teaching. *Journal of Educational Psychology*, *112*(4), 811–822. https://doi.org/10.1037/edu0000392

Fiorella, L., & Mayer, R. E. (2013). The relative benefits of learning by teaching and teaching expectancy. *Contemporary Educational Psychology*, *38*(4), 281–288. https://doi.org/10.1016/j.cedpsych.2013.06.001

Fiorella, L., & Mayer, R. E. (2014). Role of expectations and explanations in learning by teaching. *Contemporary Educational Psychology*, *39*(2), 75–85. https://doi.org/10.1016/j.cedpsych.2014.01.001

Fiorella, L., & Mayer, R. E. (2015). *Learning as a generative activity: Eight learning strategies that promote understanding*. Cambridge University Press. https://doi.org/10.1017/CBO9781107707085

Fiorella, L., & Mayer, R. E. (2016). Eight ways to promote generative learning. *Educational Psychology Review*, *28*(4), 717–741. https://doi.org/10.1007/s10648-015-9348-9

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

French, J. R. P. Jr., & Raven, B. (1959). The bases of social power. In D. Cartwright (Ed.), *Studies in social power* (pp. 150–167). Institute for Social Research, The University of Michigan.

Fukaya, T. (2013). Explanation generation, not explanation expectancy, improves metacomprehension accuracy. *Metacognition and Learning*, *8*(1), 1–18. https://doi.org/10.1007/s11409-012-9093-0

Galinsky, A. D., Gruenfeld, D. H., & Magee, J. C. (2003). From power to action. *Journal of Personality and Social Psychology*, *85*(3), 453–466. https://doi.org/10.1037/0022-3514.85.3.453

Guerrero, T. A., & Wiley, J. (2021). Expecting to teach affects learning during study of expository texts. *Journal of Educational Psychology*, *113*(7), 1281–1303. https://doi.org/10.1037/edu0000657

Guinote, A. (2007a). Power affects basic cognition: Increased attentional inhibition and flexibility. *Journal of Experimental Social Psychology, 43*(5), 685–697. https://doi.org/10.1016/j.jesp.2006.06.008

Guinote, A. (2007b). Power and goal pursuit. *Personality and Social Psychology Bulletin*, *33*(8), 1076–1087. https://doi.org/10.1177/0146167207301011

Guinote, A. (2017). How power affects people: Activating, wanting, and goal seeking. *Annual Review of Psychology*, *68*, 353–381. https://doi.org/10.1146/annurev-psych-010416-044153

Gunawardena, C. N. (1995). Social presence theory and implications for interaction and collaborative learning in computer conferences. *International Journal of Educational Telecommunications, 1*(2/3), 147–166.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.

Hayes, A. F., & Preacher, K. J. (2014). Statistical mediation analysis with a multicategorical independent variable. *British Journal of Mathematical and Statistical Psychology*, *67*(3), 451–470. https://doi.org/10.1111/bmsp.12028

Herberg, J. S., Levin, D. T., & Saylor, M. M. (2012). Social audiences can disrupt learning by teaching. *Journal of Experimental Social Psychology*, *48*(1), 213–219. https://doi.org/10.1016/j.jesp.2011.07.004

Hoogerheide, V., Loyens, S. M. M., & van Gog, T. (2014). Effects of creating video-based modeling examples on learning and transfer. *Learning and Instruction*, *33*, 108–119. https://doi.org/10.1016/j.learninstruc.2014.04.005

Hoogerheide, V., Deijkers, L., Loyens, S. M. M., Heijltjes, A., & van Gog, T. (2016). Gaining from explaining: Learning improves from explaining to fictitious others on video, not from writing to them. *Contemporary Educational Psychology*, *44–45*, 95–106. https://doi.org/10.1016/j.cedpsych.2016.02.005

Hoogerheide, V., Renkl, A., Fiorella, L., Paas, F., & van Gog, T. (2019a). Enhancing example-based learning: Teaching on video increases arousal and improves problem-solving performance. *Journal of Educational Psychology*, *111*(1), 45–56. https://doi.org/10.1037/edu0000272

Hoogerheide, V., Visee, J., Lachner, A., & van Gog, T. (2019b). Generating an instructional video as homework activity is both effective and enjoyable. *Learning and Instruction, 64*, Article 101226. https://doi.org/10.1016/j.learninstruc.2019.101226

Jacob, L., Lachner, A., & Scheiter, K. (2020). Learning by explaining orally or in written form? Text complexity matters. *Learning and Instruction, 68*, Article 101344. https://doi.org/10.1016/j.learninstruc.2020.101344

Jacob, L., Lachner, A., & Scheiter, K. (2021). Does increasing social presence enhance the effectiveness of writing explanations? *PLoS ONE, 16*(4), Article e0250406. https://doi.org/10.1371/journal.pone.0250406

Karpicke, J. D. (2017). Retrieval-based learning: A decade of progress. In J. T. Wixted (Ed.), *Cognitive psychology of memory, volume 2 of learning and memory: A comprehensive reference* (J. H. Byrne, Series ed., pp. 487–514). Academic Press. https://doi.org/10.1016/B978-0-12-809324-5.21055-9

Karpicke, J. D., Lehman, M., & Aue, W. R. (2014). Retrieval-based learning: An episodic context account. In B. H. Ross (Ed.), *The psychology of learning and motivation* (pp. 237–284). Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-800283-4.00007-1

Kearney, P., Plax, T. G., Hays, E. R., & Ivey, M. J. (1991). College teacher misbehaviors: What students don't like about what teachers say and do. *Communication Quarterly*, *39*(4), 309–324. https://doi.org/10.1080/01463379109369808

Kedrick, K., Schrater, P., & Koutstaal, W. (2023). The multifaceted role of self-generated question asking in curiosity-driven learning. *Cognitive Science, 47*(4), Article e13253. https://doi.org/10.1111/cogs.13253

Keltner, D., Gruenfeld, D. H., & Anderson, C. (2003). Power, approach, and inhibition. *Psychological Review, 110*(2), 265–284. https://doi.org/10.1037/0033-295X.110.2.265

Kipnis, D. (1972). Does power corrupt? *Journal of Personality and Social Psychology*, *24*(1), 33–41. https://doi.org/10.1037/h0033390

Kobayashi, K. (2019). Learning by preparing-to-teach and teaching: A meta-analysis. *Japanese Psychological Research*, *61*(3), 192–203. https://doi.org/10.1111/jpr.12221

Kobayashi, K. (2022). The retrieval practice hypothesis in research on learning by teaching: Current status and challenges. *Frontiers in Psychology, 13*, Article 842668. https://doi.org/10.3389/fpsyg.2022.842668

Kobayashi, K. (2023). Learning by creating teaching materials: Conceptual problems and potential solutions. *Frontiers in Psychology, 14*, Article 1095285. https://doi.org/10.3389/fpsyg.2023.1095285

Kobayashi, K. (2024). Interactive learning effects of preparing to teach and teaching: A meta-analytic approach. *Educational Psychology Review, 36*(1), Article 26. https://doi.org/10.1007/s10648-024-09871-4

Koh, A. W. L., Lee, S. C., & Lim, S. W. H. (2018). The learning benefits of teaching: A retrieval practice hypothesis. *Applied Cognitive Psychology*, *32*(3), 401–410. https://doi.org/10.1002/acp.3410

Korman, J. V., Van Quaquebeke, N., & Tröster, C. (2022). Managers are less burned-out at the top: The roles of sense of power and self-efficacy at different hierarchy levels. *Journal of Business and Psychology*, *37*(1), 151–171. https://doi.org/10.1007/s10869-021-09733-8

Kreijns, K., Xu, K., & Weidlich, J. (2022). Social presence: Conceptualization and measurement. *Educational Psychology Review*, *34*(1), 139–170. https://doi.org/10.1007/s10648-021-09623-8

Lachner, A., Ly, K.-T., & Nückles, M. (2018). Providing written or oral explanations? Differential effects of the modality of explaining on students' conceptual learning and transfer. *The Journal of Experimental Education*, *86*(3), 344–361. https://doi.org/10.1080/00220973.2017.1363691

Lachner, A., Backfisch, I., Hoogerheide, V., van Gog, T., & Renkl, A. (2020). Timing matters! Explaining between study phases enhances students' learning. *Journal of Educational Psychology*, *112*(4), 841–853. https://doi.org/10.1037/edu0000396

Lachner, A., Jacob, L., & Hoogerheide, V. (2021). Learning by writing explanations: Is explaining to a fictitious student more effective than self-explaining? *Learning and Instruction, 74*, Article 101438. https://doi.org/10.1016/j.learninstruc.2020.101438

Lachner, A., Hoogerheide, V., van Gog, T., & Renkl, A. (2022). Learning-by-teaching without audience presence or interaction: When and why does it work? *Educational Psychology Review*, *34*(2), 575–607. https://doi.org/10.1007/s10648-021-09643-4

Lammers, J., Stoker, J. I., Rink, F., & Galinsky, A. D. (2016). To have control over or to be free from others? The desire for power reflects a need for autonomy. *Personality and Social Psychology Bulletin*, *42*(4), 498–512. https://doi.org/10.1177/0146167216634064

Leung, K. C. (2019). An updated meta-analysis on the effect of peer tutoring on tutors' achievement. *School Psychology International, 40*(2), 200–214. https://doi.org/10.1177/0143034318808832

Lim, K. Y. L., Wong, S. S. H., & Lim, S. W. H. (2021). The "silent teacher": Learning by teaching via writing a verbatim teaching script. *Applied Cognitive Psychology, 35*(6), 1492–1501. https://doi.org/10.1002/acp.3881

Lim, S. W. H., Wong, S. S. H., & Visessuvanapoom, P. (2024). Durable benefits of learning-by-teaching for research question generation performance: A field experiment. *The Journal of Experimental Education*. https://doi.org/10.1080/00220973.2024.2364625. Advance online publication.

MacLeod, C., & Mathews, A. (1988). Anxiety and the allocation of attention to threat. *The Quarterly Journal of Experimental Psychology*, *40A*(4), 653–670. https://doi.org/10.1080/14640748808402292

Magni, F., Gong, Y., Li, J., Pan, J., & Zhou, M. (2024). The paradoxical relationship between sense of power and creativity: Countervailing pathways and a boundary condition. *Personnel Psychology, 77*(2), 441–474. https://doi.org/10.1111/peps.12557

Maner, J. K., Gailliot, M. T., Menzel, A. J., & Kunstman, J. W. (2012). Dispositional anxiety blocks the psychological effects of power. *Personality and Social Psychology Bulletin*, *38*(11), 1383–1395. https://doi.org/10.1177/0146167212453341

McCroskey, J. C., & Richmond, V. P. (1983). Power in the classroom I: Teacher and student perceptions. *Communication Education*, *32*(2), 175–184. https://doi.org/10.1080/03634528309378527

Meyer, B. J. (1975). *The organization of prose and its effects on memory*. North-Holland.

National Research Council. (2013). *Next generation science standards: For States, by States*. National Academies. https://doi.org/10.17226/18290

Nestojko, J. F., Bui, D. C., Kornell, N., & Bjork, E. L. (2014). Expecting to teach enhances learning and organization of knowledge in free recall of text passages. *Memory & Cognition*, *42*(7), 1038–1048. https://doi.org/10.3758/s13421-014-0416-z

Nickerson, R. S. (1999). How we know—and sometimes misjudge—what others know: Imputing one's own knowledge to others. *Psychological Bulletin, 125*(6), 737–759. https://doi.org/10.1037/0033-2909.125.6.737

Okita, S. Y., & Schwartz, D. L. (2013). Learning by teaching human pupils and teachable agents: The importance of recursive feedback. *Journal of the Learning Sciences*, *22*(3), 375–412. https://doi.org/10.1080/10508406.2013.807263

Overbeck, J. R., & Park, B. (2001). When power does not corrupt: Superior individuation processes among powerful perceivers. *Journal of Personality and Social Psychology*, *81*(4), 549–565. https://doi.org/10.1037/0022-3514.81.4.549

Pedaste, M., Mäeots, M., Siiman, L. A., de Jong, T., van Riesen, S. A. N., Kamp, E. T., Manoli, C. C., Zacharia, Z. C., & Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational Research Review*, *14*, 47–61. https://doi.org/10.1016/j.edurev.2015.02.003

Pi, Z., Guo, X., Liu, C., & Yang, J. (2025). Students with low prior knowledge learn more when explaining to a peer than to a teacher after viewing educational videos. *Active Learning in Higher Education*, *26*(1), 91–107. https://doi.org/10.1177/14697874231212260

Ribosa, J., & Duran, D. (2022). Do students learn what they teach when generating teaching materials for others? A meta-analysis through the lens of learning by teaching. *Educational Research Review, 37*, Article 100475. https://doi.org/10.1016/j.edurev.2022.100475

Ribosa, J., & Duran, D. (2023). Students' feelings of social presence when creating learning-by-teaching educational videos for a potential audience. *International Journal of Educational Research, 117*, Article 102128. https://doi.org/10.1016/j.ijer.2022.102128

Roediger, H. L. III, & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, *1*(3), 181–210. https://doi.org/10.1111/j.1745-6916.2006.00012.x

Roscoe, R. D. (2014). Self-monitoring and knowledge-building in learning by teaching. *Instructional Science*, *42*(3), 327–351. https://doi.org/10.1007/s11251-013-9283-4

Roscoe, R. D., & Chi, M. T. H. (2007). Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors' explanations and questions. *Review of Educational Research*, *77*(4), 534–574. https://doi.org/10.3102/0034654307309920

Roscoe, R. D., & Chi, M. T. H. (2008). Tutor learning: The role of explaining and responding to questions. *Instructional Science*, *36*(4), 321–350. https://doi.org/10.1007/s11251-007-9034-5

Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin*, *140*(6), 1432–1463. https://doi.org/10.1037/a0037559

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Sassenberg, K., Ellemers, N., & Scheepers, D. (2012). The attraction of social power: The influence of construing power as opportunity versus responsibility. *Journal of Experimental Social Psychology*, *48*(2), 550–555. https://doi.org/10.1016/j.jesp.2011.11.008

Schaerer, M., Tost, L. P., Huang, L., Gino, F., & Larrick, R. (2018). Advice giving: A subtle pathway to power. *Personality and Social Psychology Bulletin*, *44*(5), 746–761. https://doi.org/10.1177/0146167217746341

Scholl, A. (2020). Responsible power-holders: When and for what the powerful may assume responsibility. *Current Opinion in Psychology*, *33*, 28–32. https://doi.org/10.1016/j.copsyc.2019.06.011

Scholl, A., Ellemers, N., Scheepers, D., & Sassenberg, K. (2022). Construal of power as opportunity or responsibility. *Advances in Experimental Social Psychology*, *65*, 57–107. https://doi.org/10.1016/bs.aesp.2021.11.001

Schwartz, D. L. (1999). The productive agency that drives collaborative learning. In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 197–218). Pergamon.

Schwartz, D. L., & Lin, X. (2001). Computers, productive agency, and the effort after shared meaning. *Journal of Computing in Higher Education*, *12*(2), 3–33. https://doi.org/10.1007/BF02940954

Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications*. Wiley.

Sibley, L., Fiorella, L., & Lachner, A. (2022). It's better when I see it: Students benefit more from open-book than closed book teaching. *Applied Cognitive Psychology, 36*(6), 1347–1355. https://doi.org/10.1002/acp.4017

Sims, S., Fletcher-Wood, H., O'Mara-Eves, A., Cottingham, S., Stansfield, C., Goodrich, J., Van Herwegen, J., & Anders, J. (2025). Effective teacher professional development: New theory and a meta-analytic test. *Review of Educational Research*, *95*(2), 213–254. https://doi.org/10.3102/00346543231217480

Smith, P. K., & Hofmann, W. (2016). Power in everyday life. *Proceedings of the National Academy of Sciences, 113*(36), 10043–10048. https://doi.org/10.1073/pnas.1604820113

Smith, P. K., & Galinsky, A. D. (2010). The nonconscious nature of power: Cues and consequences. *Social and Personality Psychology Compass*, *4*(10), 918–938. https://doi.org/10.1111/j.1751-9004.2010.00300.x

Smith, P. K., Jostmann, N. B., Galinsky, A. D., & van Dijk, W. W. (2008). Lacking power impairs executive functions. *Psychological Science*, *19*(5), 441–447. https://doi.org/10.1111/j.1467-9280.2008.02107.x

Symonds, E. (2021). An 'unavoidable' dynamic? Understanding the 'traditional' learner–teacher power relationship within a higher education context. *British Journal of Sociology of Education*, *42*(7), 1070–1085. https://doi.org/10.1080/01425692.2021.1962246

Tawfik, A. A., Graesser, A., Gatewood, J., & Gishbaugher, J. (2020). Role of questions in inquiry-based instruction: Towards a design taxonomy for question-asking and implications for design. *Educational Technology Research and Development*, *68*(2), 653–678. https://doi.org/10.1007/s11423-020-09738-9

Tost, L. P. (2015). When, why, and how do powerholders "feel the power"? Examining the links between structural and psychological power and reviving the connection between power and responsibility. *Research in Organizational Behavior, 35*, 29–56. https://doi.org/10.1016/j.riob.2015.10.004

Tost, L. P., & Johnson, H. H. (2019). The prosocial side of power: How structural power over subordinates can promote social responsibility. *Organizational Behavior and Human Decision Processes*, *152*, 25–46. https://doi.org/10.1016/j.obhdp.2019.04.004

Wang, F., Cheng, M., & Mayer, R. E. (2023). Improving learning-by-teaching without audience interaction as a generative learning activity by minimizing the social presence of the audience. *Journal of Educational Psychology*, *115*(6), 783–797. https://doi.org/10.1037/edu0000801

Wittrock, M. C. (1974). Learning as a generative process. *Educational Psychologist, 11*(2), 87–95. https://doi.org/10.1080/00461527409529129

Wittwer, J., Nückles, M., Landmann, N., & Renkl, A. (2010). Can tutors be supported in giving effective explanations? *Journal of Educational Psychology*, *102*(1), 74–89. https://doi.org/10.1037/a0016727

Wong, S. S. H. (2023). Deliberate erring improves far transfer of learning more than errorless elaboration and spotting and correcting others' errors. *Educational Psychology Review, 35*(1), Article 16. https://doi.org/10.1007/s10648-023-09739-z

Wong, S. S. H. (2025). Learning by teaching with deliberate errors promotes argumentative reasoning. *Journal of Educational Psychology*, *117*(7), 1013–1038. https://doi.org/10.1037/edu0000934

Wong, S. S. H., & Lim, S. W. H. (2019). Prevention–permission–promotion: A review of approaches to errors in learning. *Educational Psychologist*, *54*(1), 1–19. https://doi.org/10.1080/00461520.2018.1501693

Wong, S. S. H., & Lim, S. W. H. (2022a). Deliberate errors promote meaningful learning. *Journal of Educational Psychology*, *114*(8), 1817–1831. https://doi.org/10.1037/edu0000720

Wong, S. S. H., & Lim, S. W. H. (2022b). The derring effect: Deliberate errors enhance learning. *Journal of Experimental Psychology: General*, *151*(1), 25–40. https://doi.org/10.1037/xge0001072

Wong, S. S. H., Lim, K. Y. L., & Lim, S. W. H. (2023). To ask better questions, teach: Learning-by-teaching enhances research question generation more than retrieval practice and concept-mapping. *Journal of Educational Psychology*, *115*(6), 798–812. https://doi.org/10.1037/edu0000802

Wong, S. S. H., & Qiu, S. X. (2026). Think first, ChatGPT later: Guiding human–AI collaboration for learning gains in independent human creativity. *Educational Psychology Review*. https://doi.org/10.1007/s10648-026-10118-7

Yap, J. B. K., & Wong, S. S. H. (2024). Deliberately making and correcting errors in mathematical problem-solving practice improves procedural transfer to more complex problems. *Journal of Educational Psychology*, *116*(7), 1112–1128. https://doi.org/10.1037/edu0000850

Authors and Affiliations

Sarah Shi Hui Wong[1,2] · **Stephen Lim**[3]

✉ Sarah Shi Hui Wong
sarahwong@smu.edu.sg

✉ Stephen Lim
Stephen.L.CogEduLab@gmail.com

[1] School of Social Sciences, Singapore Management University, 10 Canning Rise, #05-01, Singapore 179873, Singapore

[2] Division of Social Sciences, Yale-NUS College, Singapore, Singapore

[3] Cognition and Education Laboratory, Department of Educational Research and Psychology, Faculty of Education, Chulalongkorn University, 254 Phaya Thai Road, Wang Mai, Pathum Wan, Bangkok 10330, Thailand