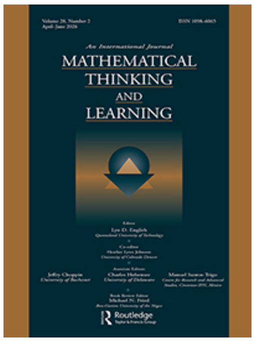

Mathematical Thinking and Learning

Routledge
Taylor & Francis Group

ISSN: 1098-6065 (Print) 1532-7833 (Online) Journal homepage: www.tandfonline.com/journals/hmtl20

Proposing an alternative direction for the development of research: a complementary perspective on Schoenfeld's approach to generality

Yusuke Uegatani & Ippo Ishibashi

MATHEMATICAL THINKING AND LEARNING
https://doi.org/10.1080/10986065.2026.2666570

Proposing an alternative direction for the development of research: a complementary perspective on Schoenfeld's approach to generality

Yusuke Uegatani[a] and Ippo Ishibashi[b]

[a]Hiroshima University High School, Fukuyama, Hiroshima, Japan; [b]Faculty of Education, Okayama University, Okayama, Japan

ABSTRACT

The purpose of this paper is to propose a theoretical framework that suggests directions for future research. While Schoenfeld's three-axis heuristic framework is well known for this purpose, it primarily points toward increasing generality. Drawing on prior studies on the generalizability of empirical findings in educational research, this paper argues that an alternative research path is possible. Building on the distinction between prevalence and scope, it proposes two types of generality: the generality of a phenomenon within a specified scope and the generality of a theory. Correspondingly, it identifies two directions for research development: delimitation of the scope and generalization of a theory. Finally, the paper argues that research development based on this framework can be understood as progressive in the Lakatosian sense. While Schoenfeld's framework suggests directions for individual studies, this framework guides competing research programmes by enabling both to progress through scope delimitation.

ARTICLE HISTORY
Received 7 August 2025
Revised 11 April 2026
Accepted 18 April 2026

KEYWORDS
Schoenfeld's heuristic framework for situating research studies; prevalence; generality; scope; delimitation of scope

Introduction

Schoenfeld's (2008) three-axis framework of generality, trustworthiness, and importance is frequently referenced when conceptualizing the nature and contributions of mathematics education studies and discussing their impacts on the field. For example, the Education Committee of the European Mathematics Society (2011) employs this framework as part of the criteria for selecting solid findings in mathematics education to present to mathematicians and teachers. Jankvist et al. (2021) employ this framework to investigate the relationships among replicability, trustworthiness, and generality in determining what should be replicated. Burkhardt and Schoenfeld (2021) argue that generality and trustworthiness are often in a trade-off relationship, depending on the sections of a research paper. Thus, it is an indispensable framework for examining the foundations of mathematics education research.

With this framework, Schoenfeld (2008) also proposed a spectrum of studies ordered by generality into categories such as limited generality (with importance), some generality, and significant generality. He referred to studies in the limited generality category, which report the existence of phenomena worth investigating, as "existence proofs" using a mathematical metaphor. This metaphor highlights that high-quality educational research cannot be fully understood or advanced by focusing solely on generality and trustworthiness. The studies in this category "may have heuristic value – they may point to issues that are important to consider, and may *turn out to be general*" (p. 500, italics added). In

CONTACT Yusuke Uegatani yuegatani@hiroshima-u.ac.jp Hiroshima University High School, Fukuyama 5-14-1, Kasuga-cho, Fukuyama-shi, Hiroshima, Japan

addition, regarding the "some generality" category, Schoenfeld discusses several studies as examples and considers how the degree of generality may be supported. For instance, he states that "the richness of the analysis lends plausibility to the generality of the findings, even though no claim is made for it" (p. 503). Similarly, in one case, he infers a connection between focal classroom observations and questionnaire data from a larger sample, stating that this "at least lends credence to the claim that the pattern of activities seen in the focal classroom, and their consequences, were anything but anomalous" (p. 503). In this way, his framework provides heuristics to enhance the quality of individual studies by suggesting directions for further research. However, it is essential to note that the framework's suggestions point only to expanding generality. We argue that there are other possible directions for improving research quality.

For example, Nesher (1987) reports that approximately 35% of Israeli sixth-grade students who had completed a unit on decimals made errors of the type "0.234 > 0.4." According to Schoenfeld's (2008) framework, this would be categorized as some generality. Note that a proportion like 35% is inherently nuanced, and in such cases, the notion of generality becomes ambiguous. Suppose we interpret this to mean that in any Israeli primary school, around 35% of students are likely to make this error. Then, the phenomenon may appear general in the sense of being widespread. However, this impression of generality may rely on an implicit assumption: namely, that in *any* Israeli primary school, an observer would encounter students who make this type of error. This does not imply that *all* Israeli primary school students are potentially predisposed to make such an error. This is just like the case where about 35% of the world's population is Chinese or Indian. That does not mean that everyone is potentially Chinese or Indian.

Schoenfeld's (2008) framework rests on the premise that "whether one is discussing quantitative or qualitative research, *generality* (or scope) and *trustworthiness* are two fundamental dimensions of research findings; and that *importance* is a third" (p. 497, italics in the original). In other words, improving research quality primarily means increasing trustworthiness and generality. It involves demonstrating how many students of a particular type are observed nationwide. However, the misconception example above suggests that high-quality research can exist even without aiming for high generality in this sense. In fact, just as the rarity of a disease does not undermine the scientific validity of medical research on rare diseases, educational research can also be scientifically valid and high-quality even when the phenomena being studied are rare.

In this regard, Ercikan and Roth's (2014) argument is particularly insightful. They argue that generalization in educational research must consider population heterogeneity and the uses of knowledge claims. For instance, if a teacher can expect that at least one student in the class has a particular type of misconception, the information is sufficiently general for instructional purposes. The same data can support different forms of generalization. For example, saying that "a large proportion of students have this misconception" is different from saying that "any classroom is likely to have a few students with this misconception." Each requires a different mode of formalization. Therefore, an alternative framework can complement Schoenfeld's (2008) framework while offering greater flexibility to accommodate diverse research in mathematics education.

The purpose of this study is to propose such an alternative heuristic framework for mathematics education research. To do so, the study focuses on identifying and separating the multiple meanings embedded in Schoenfeld's original use of the term generality. Based on this clarification, the paper proposes a new direction for research development. This direction complements the trajectory Schoenfeld initially outlined. As shown below, a clear distinction among *generality*, *prevalence*, and *scope* supports a different heuristic orientation for advancing research than the one he emphasizes.

The structure of this paper is as follows. First, we review prior studies on the generalizability of empirical findings in educational research and argue the necessity of approaching generalization with a focus on trustworthy predictions. Second, to achieve this aim, we draw on how generalization is formalized in mathematics to inform its formalization in educational research. Third, based on these insights, we define the key theoretical terms proposed in this paper. Fourth, we illustrate that the notion of *prevalence*, which was not originally part of Schoenfeld's

(2008) framework, is indeed a concern in mathematics education research. Fifth, we show that our heuristics can help individual research programmes in mathematics education become progressive in the sense of Lakatos (1978). Finally, we present our conclusions and directions for future work.

Review of prior work and positioning of the present study on generalization

Since Schoenfeld's (2008) publication, education research has advanced in examining how empirical findings can be generalized. This section relates the studies to Schoenfeld's framework and clarifies the present study's positioning within this broader context.

Prior studies on generalization in educational research

In educational research, the issue of how to generalize from a small number of cases has often been discussed in contrast to statistical generalization. For example, Ercikan (2009) examined the limitations of generalizing from a sample to a population and noted that statistical descriptions of groups do not necessarily apply to their subgroups or individual members. From phenomenological and dialectical perspectives, Roth (2009) argued that it is possible to "identify the general in the particular, the collective in the individual, or the universal in the specific" (p. 260). Drawing on Bourdieu's (1992) view of generalization, Ercikan and Roth (2014) argued that generalizability cannot be reduced to the research method alone. They emphasized that generalization requires consideration of both population heterogeneity and knowledge use.

Similar arguments have been made in the context of design research. For instance, van den Akker (2007) referred to Yin's (2018) idea of analytic generalization[1] and argued that "readers/ users need to be supported to make their own attempts to explore the potential transfer of the research findings to theoretical propositions in relation to their own context" (p. 49). More broadly, these perspectives are consistent with Cobb's (2007) pragmatist view, which draws on Dewey and holds that "*the truth of fallible, potentially revisable ideas is justified primarily in terms of the insight and understanding they give into learning processes and the means of supporting their realization*" (p. 14, italics in the original).

Despite differences in philosophical orientation, the theoretical view of generalization represented by scholars such as Cobb (2007) and Ercikan and Roth (2014) blurs the boundary between qualitative and quantitative approaches. This idea has recently been further refined by Foster (2023), who attempts to bridge the divide between qualitative and quantitative researchers by focusing on whether research is exploratory or confirmatory.

Interpreting prior studies through the lens of Schoenfeld's framework

As discussed above, theoretical discussions about the generalizability of educational research findings have deepened in recent years. However, these discussions diverge from the notion of generality as proposed in Schoenfeld's (2008) framework. From Schoenfeld's perspective, prior theoretical discussions may appear to focus solely on generalization, while neglecting the dimension of trustworthiness. As Foster (2023) rightly observes, individual research papers often make generalized claims in their discussion sections that are unsupported by the presented data (see also Burkhardt and Schoenfeld's (2021) discussion). Even for nearly universal claims, Schoenfeld (2008) warns that whether they apply in specific situations should be approached with caution: "caveat emptor is probably the best attitude" (p. 504). In his framework, the context of knowledge use is primarily discussed in terms of trustworthiness rather than generalization.

Interestingly, Dawkins and Karunakaran (2016) point out a related issue: Some case studies are too theory-laden. That means that theory-laden insights from a particular case do not necessarily ensure trustworthy predictions of future cases.

Positioning of the present study

While prior studies on the generalizability of empirical findings consider population heterogeneity and knowledge use, Schoenfeld (2008) offers a valuable insight into generalizability and trustworthy predictions. Therefore, integrating these discussions and exploring an alternative framework that highlights directions of research not visible through Schoenfeld's (2008) framework can contribute to a more inclusive understanding of diverse approaches in mathematics education research.

Insights for formalization from mathematical metaphors

As discussed in the introduction of this paper, the direction of generalization is closely related to its formalization. For the following two reasons, this section seeks insight into educational research from how generalization is formalized in mathematics. First, educational research, including the work of Schoenfeld (2008), has long drawn on mathematical metaphors when examining its argumentative grammar (e.g., Bakker, 2018; Lampert, 1990; Schoenfeld, 2008). The value of this practice has not diminished, and the formalization of generalization in mathematics should provide an important clue for developing perspectives that complement Schoenfeld's (2008) approach. Second, as mathematics education researchers are familiar with mathematics as a discipline, drawing on its formalization of generalization is likely to yield formulations that are more accessible and intuitive to researchers in this field.

Consider Euler's prime-generating polynomial $n^2 - n + 41$ as an example. The generality of the statement "For natural numbers n, $n^2 - n + 41$ is prime" is typically assessed in mathematical terms by whether the claim holds for all natural numbers. However, this claim fails for $n = 41$, providing a counterexample. Presenting $n = 41$ thus serves as an existence proof for the statement, "There exists a natural number for which $n^2 - n + 41$ is not prime." As this example illustrates, generality does not form a spectrum in mathematical terms.

On the other hand, Schoenfeld's (2008) concept of generality differs in meaning. He defines generality as follows: "*Generality, or Scope*. The *claimed generality* of a study is the set of circumstances in which the author(s) of a study claim that the findings of the study apply. The potential generality of a study is the set of circumstances in which the results of the study (if trustworthy) might reasonably be expected to apply" (p. 497, italics in the original). Applying this notion of generality to Euler's prime-generating polynomial, the *generality (scope)* would refer to the range $1 \leq n \leq 40$.

Schoenfeld's (2008) use of generality has some merit. For instance, if it were argued that $n^2 - n + 41$ is prime for $n = 1, 2, 3$, this claim could serve as an existence proof of a phenomenon with limited generality. Similarly, if this claim were extended to hold for $n = 4, 5, 6, \cdots\cdots, 40$, it could be said to exhibit some generality. However, the claim that $n^2 - n + 41$ is prime does not achieve significant generality in Schoenfeld's (2008) sense. For instance, while 86% of values are prime within the range $1 \leq n \leq 100$, this proportion drops to 58.1% for $1 \leq n \leq 1000$. We do not usually call these statements general in mathematics.

Introducing theoretical terms

To enable a more rigorous discussion, we propose introducing the concepts of prevalence and scope, defined as follows:

Prevalence: The proportion of elements within a specified set (or domain) that possess a particular characteristic.

Scope: The specified set (or domain) itself.

Let us illustrate the distinction between prevalence and scope using the example of the prime-generating formula mentioned in the previous section. For instance, the statement "$n^2 - n + 41$ yields 86% primes within $1 \leq n \leq 100$, but only 58.1% within $1 \leq n \leq 1000$" holds. In this case, the percentages 86% and 58.1% represent the prevalence, while the ranges $1 \leq n \leq 100$ and $1 \leq n \leq 1000$ define the scope.

Similarly, in mathematics education research, it is also possible to distinguish between *prevalence* and *scope*. For example, in Nesher's (1987) study, the scope refers to Israeli sixth-grade students who had completed the unit on decimals, and the prevalence of the misconception is approximately 35%. However, this might also be interpreted as follows: the scope refers to Israeli primary schools, and the prevalence of schools with at least one student who holds the misconception is close to 100%. Even when researchers believe they are focusing on the same phenomenon, whether it appears general or not can depend heavily on how the scope is defined.

This means that the establishment of generality depends not on expanding the scope but on defining a scope in which prevalence is sufficiently high. That is, generality is not secured through *generalization*—in the sense of widening the scope – but through *delimitation*—the act of narrowing the scope appropriately.

For example, Kim et al. (2012) suggest, through a comparative study of English-speaking and Korean-speaking university students, that the concept of "infinity" is processual for the former, and that its objectification poses particular difficulties. They enhance the prevalence of this phenomenon by limiting the scope to English-speaking students, whereas it would not appear prevalent if the scope were defined as university students worldwide.

Such delimitation corresponds to what Schoenfeld (2008) refers to as the *specification of applicability conditions*. When presenting various theories as examples of research with significant generality, he notes: "various theories [. . .] all have their applicability conditions. It is the responsibility of theorists to specify those conditions, to define the relevant constructs, and to address the limits (as well as the strengths) of what the theories can actually explain" (p. 505).

Viewed in this way, it becomes clear that *establishing generality* is easy to achieve. If one refrains from any form of generalization and simply restricts the scope to a single, specific learner, then any claim about that learner is trivially general within that scope. However, we typically do not assign much value to such trivial forms of generality. There are degrees of values of generality: Some are more valuable than others.

A key to approaching this issue is the role of theory. If, within a given scope, a phenomenon is not only empirically prevalent but also accompanied by a theoretical account explaining *why* it is prevalent in that scope, then the theory gains in value. In this sense, we should be concerned with the extent to which each study contributes to the development of scientific theory in mathematics education.

With this in mind, we propose the following use of terms:

Generality of a phenomenon (within a specified scope): High prevalence of the phenomenon within the scope.

Generality of a theory: The breadth of the scope to which the theory applies.

Delimitation of the scope (within which a phenomenon is or will be observed): The act of narrowing the scope to which a theory applies, thereby establishing the generality of the phenomenon and providing a theoretical account of *why* the phenomenon is or will be prevalent within that scope.

Generalization of a theory (or of the scope within which a phenomenon is or will be observed): The act of expanding the scope to which the theory applies, while providing a theoretical account of *why* the phenomenon remains prevalent within the broader scope.

As noted in Schoenfeld's (2008) discussion of Category 3, *significant generality* does not necessarily imply *universality*. In educational research, achieving 100% prevalence is rarely possible. Therefore, when considering the generality of a phenomenon, it is reasonable to treat the required level of prevalence as context-dependent.

On the other hand, when shifting our focus from individual research studies to theory, we can compare two theories and consider which is more general. For example, we may say that Tall's (2011) concept of the *crystalline concept* is more *general* than Gray and Tall's (1994) notion of the *procept*. The latter was proposed to account for the flexible use and interpretation of symbols, particularly in the context of algebra and analysis. In contrast, the former is a more abstract concept intended to apply across the entire domain of mathematics. This is an example of a discussion not about the generality of a phenomenon, but about the relative generality of the two theories.

By distinguishing between the generality of a phenomenon and the generality of a theory in this way, we can also clearly differentiate the two opposing research processes of *delimitation* and *generalization*. The former is associated with the establishment of the generality of a phenomenon, while the latter is concerned with expanding the generality of a theory.

These distinctions enable us to express more nuanced forms of generality. They are also helpful in the context of statistical inference. For example, when estimating the prevalence of a phenomenon in a population based on its prevalence in a sample, we are concerned with the generality of prevalence, that is, whether the prevalence observed within the sample can be expected to hold within the broader scope (i.e., the population).

This section can be summarized as follows. In Schoenfeld's (2008) framework, generality and scope are not clearly distinguished. By contrast, the formalization of generalization in mathematics allows for clear distinctions among generality, prevalence, and scope. Drawing on this distinction, we proposed that these terms should also be clearly distinguished and defined in mathematics education research. Based on these definitions, we distinguished four notions: the generality of a phenomenon, the generality of a theory, the delimitation of the scope, and the generalization of a theory. This distinction allows identification of research directions that are not necessarily suggested by Schoenfeld's (2008) framework alone. When considering issues of generality, distinguishing among generality, prevalence, and scope helps clarify future directions for inquiry. In the remainder of the paper, we use these terms in the sense defined in this section.

Examples: interest in prevalence in mathematics education research

Based on the definitions presented in the previous section, let us consider the following two groups of studies in mathematics education as examples: A group of quantitative studies on preservice teachers' conceptions of multidigit whole numbers and a group of case studies on university students' conceptions of integration. In the previous section, we provided examples only to the extent necessary to explain the theoretical terms. In this section, however, we aim to demonstrate, through new examples, how the concepts of *prevalence* and *generality* can become intricately intertwined in actual mathematics education research. We also aim to demonstrate how the newly defined theoretical terms can serve as a heuristic framework for facilitating the advancement of individual research studies to their following stages.

Studies on preservice teachers' conceptions of multidigit whole numbers

The following two quantitative studies on preservice teachers' conceptions of multidigit whole numbers are excellent examples of quantitative research focusing on *prevalence* in mathematics education research: the study by Thanheiser (2010) and its replication study by Jacobson and Simpson (2019). In the original study by Thanheiser (2010), 9% of preservice teachers had correct conceptions (95% CI [0.007, 0.189]), while in the replication study by Jacobson and Simpson (2019), the percentage was 18% (95% CI [0.093, 0.261]). These studies primarily focus on prevalence, and it is evident that their research quality is not diminished merely because the prevalence rates are 9% or

18%. In addition, this prevalence may change with improvements in teacher education, which is not a research quality issue.

On the other hand, these studies perform interval estimation. While we will not discuss what interval estimation might imply in nonrandom sampling situations, it is clear that these studies are concerned with how well the prevalence observed in the sample survey holds in the population. This reflects a focus on whether the prevalence obtained from the sample survey can be generalized to the prevalence in the population, aligning with our concept of generality. In short, the issue here is *the generality of the prevalence*. Furthermore, an interest in the generality of the prevalence essentially entails a shift in focus from individual groups of prospective teachers to the set of such groups. A single study can address both prevalence and generality; we should, thus, distinguish between the three concepts of generality, prevalence, and scope.

Our newly proposed theoretical terms function as heuristic frameworks for further developing these research studies. Jacobson and Simpson's (2019) replication study successfully replicated Thanheiser's (2010) original study in terms of the proportion of correct conceptions and the most prevalent sub-conception. However, they also observed differences in the distribution of participants' responses across the tasks between the replication and the original studies. They pointed to the difference in the mathematics coursework taken by the participants as a possible explanatory factor.

If the researchers were to succeed in providing a theoretical account of how each type of coursework gave rise to the observed response distributions, then this proposed explanatory factor would constitute a *delimitation* of the phenomenon. Since both the original and replication studies were conducted in the United States, if subsequent replication studies in the U.S. found that groups of prospective teachers who had taken particular mathematics courses exhibited similar response distributions with high prevalence, the delimitation would be considered successful.

Furthermore, it would also be possible to *generalize* the theory beyond the U.S. and to test *the generality of the theory* in other countries. Of course, if the replication studies reveal that the hypothesized factor is unlikely to account for the observed differences, then the research would return to the stage of formulating a new hypothesis.

Case studies on university students' conceptions of integration

The following two case studies on conceptions of integration serve as examples of how interest in prevalence evolves in non-quantitative mathematics education research: Stevens and Jones (2023), which investigated university students' meanings of integration in interview settings, and Lehmann (2026), which conducted a case-based examination to determine whether the phenomena observed in interview settings could also be observed in classroom environments. While Lehmann (2026) noted that most findings were consistent with prior research, he argued that "the contributions of this study lie in the contrasting findings" (p. 24) and highlighted two aspects that did not align with those reported by Stevens and Jones (2023). First, while Stevens and Jones (2023) reported in their pilot study that introducing a graphical representation too early often led students to lose sight of the quantities, Lehmann (2026) suggests that the teacher/researcher's suggestion of the trapezoidal rule helped students retain their focus on the quantities. Specifically, in Stevens and Jones (2023), university students interviewed in their study lost sight of the quantities or struggled to formulate the target quantity because a graphical representation was introduced prematurely. In contrast, Lehmann (2026) suggests that high school students in the observed class avoided such difficulties due to the teacher/researcher's interventions and instructional design, which were based on the trapezoidal rule. This can be understood as an attempt to explore the conditions under which a particular phenomenon occurs (or does not occur).

While individual design studies may fall into the category of limited generality under Schoenfeld's (2008) framework, they can be seen as an interventionist approach to investigating the conditions under which a given phenomenon becomes highly prevalent. Based on our definition, design research aims to enhance *prevalence* by appropriately *delimiting the scope*, thereby establishing the *generality* of the phenomenon. Indeed, design research directly investigates the validity of theoretical accounts that explain why a given phenomenon occurs under particular conditions.

Interestingly, Schoenfeld (2008) noted, "large *n* is no guarantee of either trustworthiness or generality" (p. 499, italics in the original). Similarly, we can say that if design research clearly specifies its scope and successfully reproduces the expected outcomes, it can make a meaningful contribution to establishing the generality of the phenomenon. This holds even with a small sample size, provided that the sample is randomly drawn from the defined scope. Thus, there is no need to assume that design research has inherently limited generality. However, random sampling is likely to be unrealistic in design research. In practice, the declared scope is often overly broad, and the expected outcomes may not be successfully reproduced through design research. In such cases, we are likely to engage in repeated delimitation to refine the scope appropriately. Of course, if the expected outcomes continue to be successfully reproduced, it would then be reasonable to consider generalizing the theory and applying it to a broader scope.

Compatibility with the Lakatosian theory of research programmes

The heuristic framework proposed in this paper is highly compatible with Lakatos's (1978) theory of research programmes. In this section, we examine how our heuristic framework can support the development of a progressive research programme in mathematics education in the Lakatosian sense. Schoenfeld (2008) referred to Kuhn (1970) in addressing the philosophy of science, stating only that "educational research has hardly entered a period of 'normal science'" (Schoenfeld, 2008, p. 468). In contrast, this paper introduces a Lakatosian perspective to show that the proposed framework supports the healthy development of educational research from the standpoint of the philosophy of science. In particular, introducing this Lakatosian perspective contributes to establishing a research direction that accounts for the context of knowledge use while maintaining a commitment to testing the trustworthiness of predictions.

According to Lakatos (1978), a research programme includes two types of hypotheses that structure the theory it supports: the *hard core*, which consists of the central assumptions, and the *protective belt*, which consists of auxiliary hypotheses. Even when a theory within a research programme faces empirical refutation, the programme need not be abandoned entirely. Instead, it may discard or revise hypotheses in the protective belt to preserve the hard core. Based on how the programme responds to such refutations, it is classified as either progressive or degenerative. A progressive research programme is one in which the reconstructed theory not only explains the previous counterexamples but also successfully predicts new empirical data. In contrast, a degenerative programme lacks such explanatory and predictive power. Only the former can avoid falling into pseudoscience.

While Schoenfeld's (2008) original framework provides valuable guidance for how a single study might proceed, it does not address how competing research programmes should develop in relation to one another. For instance, consider a case where one study makes a trustworthy claim that a specific phenomenon demonstrates significant generality, and a subsequent study provides an existence proof of a counterexample. Schoenfeld's framework, which focuses on increasing generality, offers no heuristics for resolving the conflict that arises between these two studies.

In contrast, our heuristic framework supports the development of a progressive research programme. Let us illustrate this through the relationship between Piaget-based constructivism (cf. von Glasersfeld, 1995) and sociocultural theories in mathematics education (cf. Lerman, 2000). Cobb et al. (1992) aimed to avoid the misinterpretation of Piaget-based constructivism as suggesting that teachers should not teach students anything. They proposed a view that bridges the dualism between mathematics in students' heads and mathematics in their environment. This view sought to integrate Piaget-based and sociocultural perspectives in a complementary manner (see also Cobb, 1994). On the other hand, the philosophical shift known as the "social turn" (Lerman, 2000) led to a broader trend of developing Vygotsky-based theories without Piaget-based ideas (e.g., Ernest, 1998; Radford, 2021). While Confrey and Kazak (2006) argue that constructivism remains progressive in the Lakatosian sense, they also pessimistically remark that "constructivism's influence on mathematics education is, in our opinion, unfortunately, waning" (p. 331).

Our framework, however, suggests that such a pessimistic view may not be necessary. *Delimiting the scope* is essential for constructivism to make trustworthy predictions about future empirical data. Before the social turn, there was an overly broad expectation that constructivism and its associated bridging theories could predict all forms of mathematics learning across contexts. This expectation was indeed too general. Yet, as demonstrated by Simon et al. (2010, 2018a, 2018b), mathematical activities based on theoretically designed task sequences can support conceptual development. Through such teaching experiments with individual students, they showed that constructivism and its bridging theories can still predict future empirical data within clearly defined conditions. Because such theories cannot predict the complexity of social interactions between researchers and students, Simon and colleagues excluded data from teaching experiments in which unexpected social interactions occurred during testing. These examples illustrate how to appropriately delimit the scope to establish the generality of a phenomenon in mathematics education research.

In today's context, it may be less productive to focus on which "−ism" dominates the mathematics education community. Even scholars who align with a Vygotskian perspective may at times find constructivism and its bridging theories useful for predicting conceptual development in teaching experiments with limited social interaction. Rather than evaluating research programmes by the number of supporters they have, it may be more fruitful to consider whether they continue to generate progressive developments. Programmes with a narrow scope can still be progressive, and their value should not be dismissed merely because of their limited scope. Ensuring that such progressive programmes are sustained may therefore be an important consideration.

In sum, our framework suggests delimiting the scope as a possible path forward for competing research programmes. Each programme can redefine its scope so that the phenomenon of interest occurs with high prevalence within that scope. This enables both programmes to make trustworthy predictions about future empirical data and remain progressive. Such an approach not only responds to concerns about the use of knowledge, emphasized in earlier work on the generalizability of empirical findings by Cobb (2007) and Ercikan and Roth (2014), but also fulfills the emphasis on trustworthiness seen in Schoenfeld (2008) and Foster (2023).

Conclusion

This paper proposes an alternative framework for suggesting a research direction different from the one Schoenfeld outlined. We first distinguished between *prevalence* and *scope*, and then argued that generality should be considered in two forms: the *generality of a phenomenon* (within a specified scope) and the *generality of a theory*. Correspondingly, we identified two directions in which research studies may develop: *delimiting the scope and generalizing a theory (or its scope)*.

By drawing on a Lakatosian perspective, our framework also suggests how research programmes can become progressive. *Delimiting the scopes* offers promising paths for competing programmes to move forward. Moreover, for a research programme to be progressive, the focus should not be solely on generalizing claims but also on enhancing the trustworthiness of predictions. Such an approach contributes to a commitment both to the context of knowledge use, which is emphasized in prior studies on the generalization of empirical findings in education by Cobb (2007) and Ercikan and Roth (2014), and to *trustworthiness*, which was highlighted by Schoenfeld (2008) and Foster (2023).

While this paper has provided an alternative path for research development to Schoenfeld's suggestion, it has not addressed the other two dimensions he proposed: aspects of *trustworthiness* beyond prediction and *importance*. Consideration of these dimensions should remain grounded in Schoenfeld's original framework. Addressing how to improve them remains a task for future work.

Note

1. We cite a more recent edition of Yin than the ones referred to by van den Akker (2007) and Dawkins and Karunakaran (2016).

Acknowledgments

In accordance with the Taylor & Francis AI Policy, the authors declare that the following generative Artificial Intelligence (AI) tools were used solely for language improvement and not for idea generation or literature review: ChatGPT 5, Grammarly (ver. 1.2), and Microsoft Editor for Word 365 MSO (version 2510). All opinions and views expressed in this paper are solely those of the authors and do not necessarily reflect the views of the JSPS and the AI tools.

Authors' contribution statement

Both authors contributed to the literature review. The first author wrote the first draft of the manuscript, which both authors commented on. Both authors read and approved the final manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Japan Society for the Promotion of Science (Grant Numbers: JP24K00422 and JP23K12847).

Notes on contributors

Yusuke Uegatani, who holds a PhD, is a mathematics teacher at Hiroshima University High School, Fukuyama, one of the affiliated research schools of Hiroshima University. His research interests fall into two main areas. The first concerns the analysis of mathematical discourse and the design of classroom instruction grounded in inferentialism. In this area, he conducts transdisciplinary research in collaboration with scholars in inferentialist philosophy, philosophy of education, and mathematics education. The second focuses on theorizing the relationship between practice and theory. As a teacher, he draws on theories of mathematics education in his daily practice, while as a researcher, he feeds empirical classroom data back into theory development. He views this reciprocal process itself as a worthwhile object of study.

Ippo Ishibashi is a Senior Assistant Professor in Mathematics Education at Okayama University. He holds a PhD in Education. His research interests can be broadly divided into two areas. One area concerns students' responses that are commonly interpreted as typical incorrect answers. In this work, he investigates the reasons behind such responses and develops lesson designs intended to help students move toward more desirable forms of reasoning, with a particular focus on probability and word problem solving. The other area addresses the relationship between educational practice and theory. Through this work, he seeks to clarify how teachers and schools, the academic community, and policymakers can work together to support children's learning.

ORCID

Yusuke Uegatani http://orcid.org/0000-0002-6142-6078
Ippo Ishibashi http://orcid.org/0000-0002-3086-1608

References

Bakker, A. (2018). *Design research in education: A practical guide for early career researchers*. Routledge. https://doi.org/10.4324/9780203701010

Bourdieu, P. (1992). The practice of reflexive sociology (the Paris workshop). In P. Bourdieu & L. J. D. Wacquant (Eds.), *An invitation to reflexive sociology* (pp. 216–260). University of Chicago Press.

Burkhardt, H., & Schoenfeld, A. (2021). Not just "implementation": The synergy of research and practice in an engineering research approach to educational design and development. *ZDM - Mathematics Education*, *53*(5), 991–1005. https://doi.org/10.1007/s11858-020-01208-z

Cobb, P. (1994). Where is the mind? Constructivist and sociocultural perspectives on mathematical development. *Educational Researcher*, *23*(7), 13–20. https://doi.org/10.3102/0013189X023007013

Cobb, P. (2007). Putting philosophy to work: Coping with multiple theoretical perspectives. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics* (Vol. 1, pp. 3–38). Information Age Publishing.

Cobb, P., Yackel, E., & Wood, T. (1992). A constructivist alternative to the representational view of mind in mathematics education. *Journal for Research in Mathematics Education*, *23*(1), 2–33. https://doi.org/10.2307/749161

Confrey, J., & Kazak, S. (2006). A thirty-year reflection on constructivism in mathematics education in PME. In A. Gutiérrez & P. Boero (Eds.), *Handbook of research on the psychology of mathematics education: Past, present and future* (pp. 305–345). Sense Publishers.

Dawkins, P. C., & Karunakaran, S. S. (2016). Why research on proof-oriented mathematical behavior should attend to the role of particular mathematical content. *The Journal of Mathematical Behavior*, *44*, 65–75. https://doi.org/10.1016/j.jmathb.2016.10.003

Education Committee of the European Mathematics Society. (2011). "Solid findings" in mathematics education. *Newsletter of the European Mathematical Society*, *81*, 46–48.

Ercikan, K. (2009). Limitations in sample-to-population generalizing. In K. Ercikan & W.-M. Roth (Eds.), *Generalizing from educational research: Beyond qualitative and quantitative polarization* (pp. 211–234). Routledge.

Ercikan, K., & Roth, W.-M. (2014). Limits of generalizing in education research: Why criteria for research generalization should include population heterogeneity and uses of knowledge claims. *Teachers College Record*, *116*(4), 1–28. https://doi.org/10.1177/016146811411600405

Ernest, P. (1998). *Social constructivism as a philosophy of mathematics*. SUNY Press.

Foster, C. (2023). Methodological pragmatism in educational research: From qualitative-quantitative to exploratory-confirmatory distinctions. *International Journal of Research & Method in Education*, *47*(1), 4–19. https://doi.org/10.1080/1743727X.2023.2210063

Gray, E. M., & Tall, D. O. (1994). Duality, ambiguity, and flexibility: A "proceptual" view of simple arithmetic. *Journal for Research in Mathematics Education*, *25*(2), 116–140. https://doi.org/10.5951/jresematheduc.25.2.0116

Jacobson, E., & Simpson, A. (2019). Prospective elementary teachers' conceptions of multidigit number: Exemplifying a replication framework for mathematics education. *Mathematics Education Research Journal*, *31*(1), 67–88. https://doi.org/10.1007/s13394-018-0242-x

Jankvist, U. T., Aguilar, M. S., Misfeldt, M., & Koichu, B. (2021). What to replicate? *Implementation and Replication Studies in Mathematics Education*, *1*(2), 141–153. https://doi.org/10.1163/26670127-01010015

Kim, D.-J., Ferrini-Mundy, J., & Sfard, A. (2012). How does language impact the learning of mathematics? Comparison of English and Korean speaking university students' discourses on infinity. *International Journal of Educational Research*, *51-52*, 86–108. https://doi.org/10.1016/j.ijer.2012.01.004

Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). University of Chicago Press.

Lakatos, I. (J. Worrall & G. Currie (Eds.). (1978). *The methodology of scientific research programmes* (Vol. 1). Cambridge University Press.

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Educational Research Journal*, *27*(1), 29–63. https://doi.org/10.3102/00028312027001029

Lehmann, T. H. (2026). Sustaining quantitatively-grounded meanings for definite integrals in high school calculus. *Mathematical Thinking and Learning*, *28*, 23–52. https://doi.org/10.1080/10986065.2024.2384689

Lerman, S. (2000). The social turn in mathematics education research. In J. Boaler (Ed.), *Multiple perspectives on mathematics teaching and learning* (pp. 19–44). Ablex.

Nesher, P. (1987). Towards an instructional theory: The role of student's misconceptions. *For the Learning of Mathematics*, *7*(3), 33–40.

Radford, L. (2021). *The theory of objectification: A Vygotskian perspective on knowing and becoming in mathematics teaching and learning*. Brill.

Roth, W.-M. (2009). Phenomenological and dialectical perspectives on the relation between the general and the particular. In K. Ercikan & W.-M. Roth (Eds.), *Generalizing from educational research: Beyond qualitative and quantitative polarization* (pp. 235–260). Routledge.

Schoenfeld, A. H. (2008). Research methods in (mathematics) education. In L. English (Ed.), *Handbook of international research in mathematics education* (2nd ed. pp. 467–519). Routledge.

Simon, M. A. (2018a). An emerging methodology for studying mathematics concept learning and instructional design. *The Journal of Mathematical Behavior*, *52*, 113–121. https://doi.org/10.1016/j.jmathb.2018.03.005

Simon, M. A., Kara, M., Placa, N., & Avitzur, A. (2018b). Towards an integrated theory of mathematics conceptual learning and instructional design: The learning through activity theoretical framework. *The Journal of Mathematical Behavior*, *52*, 95–112. https://doi.org/10.1016/j.jmathb.2018.04.002

Simon, M., Saldanha, L., McClintock, E., Akar, G. K., Watanabe, T., & Zembat, I. O. (2010). A developing approach to studying students' learning through their mathematical activity. *Cognition & Instruction*, *28*(1), 70–112. https://doi.org/10.1080/07370000903430566

Stevens, B. N., & Jones, S. R. (2023). Learning integrals based on adding up pieces across a unit on integration. *International Journal of Research in Undergraduate Mathematics Education*, *9*(1), 118–148. https://doi.org/10.1007/s40753-022-00204-w

Tall, D. (2011). Crystalline concepts in long-term mathematical invention and discovery. *For the Learning of Mathematics*, *31*(1), 3–8.

Thanheiser, E. (2010). Investigating further preservice teachers' conceptions of multidigit whole numbers: Refining a framework. *Educational Studies in Mathematics*, *75*(3), 241–251. https://doi.org/10.1007/s10649-010-9252-7

van den Akker, J. (2007). Curriculum design research. In T. Plomp & N. Nieveen (Eds.). An introduction to educational design research: Proceedings of the seminar conducted at the East China Normal University, November 23-26, 2007, Shanghai, China (pp. 37–50). Netherlands Institute for Curriculum Development.

von Glasersfeld, E. (1995). *Radical constructivism: A way of knowing and learning*. The Falmer Press.

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE Publications.