

Teachers' Expectation About Geometric Calculations in High School Geometry

Nicolas Boileau
University of Michigan
nboilea@umich.edu

Patricio Herbst
University of Michigan
pgherbst@umich.edu

This paper reports on a study of the instructional situation in high school Geometry that Hsu (2010) called Geometric Calculation in Algebra (GCA). In particular, we conducted a virtual breaching experiment in order to examine the extent to which high school teachers recognized breaches of two norms that we conjectured to describe geometry teachers' expectations of this work context. The results of our analysis of the data (using z-tests and mixed effect regression models) provide evidence that, in the situation of GCA, (1) teachers appear not to take issue with giving students tasks that require them to set-up and solve equations whose solutions have no geometric meaning (e.g., the length of a side of the figure is zero), and (2) teachers do not appear to expect students to document the geometric theorem or property that justify the setup of those equations (highlighting the contrast between the situation GCA and that of doing proofs).

Keywords: Algebra and Algebraic Thinking, Geometry and Geometrical and Spatial Thinking, Instructional Activities and Practices, Measurement

In an effort to describe the practice of teaching mathematics in schools, some mathematics educators have argued for the importance of understanding the factors that influence mathematics teachers' instructional decisions. Some researchers have described these decisions as motivated by the teacher's individual goals, beliefs, and knowledge (Schoenfeld, 2010). However, others have sought to complement this perspective by suggesting that decisions might also be understood as influenced by customary norms of the practice of teaching. In their account of the practical rationality of mathematics teaching, Herbst and Chazan (2012) suggest that students and teachers recognize some patterns of interaction as defaults for recurring classroom situations (e.g., when doing proofs in high school geometry class).

This study investigates norms of another instructional situation in high school geometry instruction: what Hsu (2010) has referred to as Geometric Calculation in Algebra. We use Herbst's (2006) definition of instructional situation to conceptualize geometric calculations in algebra (GCA): GCA enables an exchange between the work students do posing and solving an equation and the claim their teacher can make that they know properties about a geometric figure (to which the equation refers). As in the case of other instructional situations, we expect there are norms for what the teacher and the student are expected to do to enable them to do such work and operate such exchange. Figure 1 shows an example of a GCA task.

1. Determine the length of each of the sides of the isosceles triangle, below.

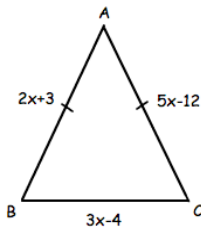


Figure 1: Sample GCA task

The example in Figure 1 illustrates some norms of GCA: Students are informed (often, through a diagram) that certain dimensions of a given geometric figure can be represented by given algebraic expressions and they are expected to use their knowledge of the properties of that figure to set up and solve algebraic equations using the given expressions in order to find one or more of the figure's dimensions. Their success in this work counts toward their understanding of the geometric properties of the figure as well as their retention of algebraic skill.

The work of investigating the norms of this instructional situation may be of interest to mathematics educators (both practitioners and researchers) for several reasons. For one, in terms of efforts to understand and describe the ways that mathematics is actually taught and learned in schools, norms provide interesting insights because they represent both expected actions and a rationale for teacher's instructional actions. Norms also provide a baseline in the process of improving instruction: Efforts to change instruction need to oppose existing norms and propose justifiable breaches of such norms. Along those lines, while this study explores GCA in particular, its methods are equally applicable to the study of norms of other situations in other courses of study beyond high school geometry. Second, for those particularly interested in geometry instruction, the situation of GCA is a common instructional situations in American high school geometry courses, and one that provides students with opportunities to engage in practices that have generally been supported by mathematics educators, such as engaging in algebraic and geometric reasoning as well as connecting multiple representations (NCTM, 2014). Further, Hsu (2010) argued that geometric calculation tasks offer students inroads to the types of reasoning needed for understanding and writing proofs in geometry.

Based on our observations of American high school geometry classrooms and informal analysis of geometry textbooks, we hypothesized that the following were norms of the instructional situation of GCA:

1. When a GCA task is given to students the algebraic expressions associated to the dimensions of the figure are such that when an equation is set-up on the basis of one or more true geometric properties of the figure the numerical measures obtained from the solution of such equation will have interpretable geometric meanings (e.g., side lengths and angle measures will be positive).
2. Although students may be asked to state orally the geometric property that they use to set-up one or more equations when solving GCA problems at the board, they are not expected to write that property.

For sake of brevity, we will refer to the first of these two norms as the GCA Figure (GCAF) norm and to the second as the GCA Theorem (GCAT) norm.

Similar to other mathematics educators who have endeavored to investigate instructional norms (e.g., Dimmel, 2015; Herbst, Kosko, & Dimmel, 2013), we adopted a variation of a breaching experiment (Garfinkel, 1963) to determine whether two norms that we conjectured to exist actually describe how high school geometry teachers expect work on GCA problems will unfold, by examining the extent to which high school mathematics teachers recognize breaches of them. Accordingly, we posed the following two research questions:

1. Do the GCAF and GCAT norms exist (i.e., represent how high school geometry teachers expect work on GCA problems will unfold)?
2. How do participants react to breaches of the GCAF and GCAT norms?

Further, aware that in any instance of an instructional situation more than one norm might be breached, we also sought to investigate how breaches of a given norm at one point in a lesson

might influence teacher's reactions to breaches of norms of that situation that occur later in the lesson. We therefore also posed a third research question:

3. Is participants' recognition of breaches of the GCAT norm affected by whether the GCAF norm is breached or complied with?

In order to measure teachers' recognition of breaches of these hypothesized norms, we designed and implemented a research instrument, which we describe in the next section.

Methods

Data collection

The present study represents a first attempt at answering our three research questions – by designing a research instrument and using it with a convenience sample of 40 high school mathematics teachers from a Midwestern state. The instrument designed sought to elicit mathematics teachers' reactions to storyboard representations of classroom scenarios, through online multimedia questionnaires, in order to determine whether hypothesized norms of the instructional situation of GCA exist, by measuring the extent to which participants recognize the breaches of those hypothesized norms (see Herbst, Kosko, and Dimmel, 2013) that occurred in some of those scenarios.

The instrument was comprised of twelve item sets. Each item set consisted of various questions referring to one storyboard. The scenario represented by each storyboard can be described as following one of three experimental conditions, which we refer to as the CFCT, the CFBT and the BFBT conditions (to denote which of the two norms are breached or complied with). The four CFCT scenarios were conjectured to be completely normative (both norms were complied with). The four CFBT scenarios breach the GCAT norm, but not the GCAF norm. The four BTBF scenarios breach both norms. We represented four scenarios of each condition to increase the “construct validity” (Shadish, Cook, & Campbell, 2002) of the instrument – only having one scenario of each condition would threaten the validity of our claim that participants' responses to the scenarios represent whether and how participants might react to breaches or compliance with the two norms (rather than the specific task or other incidental aspects of the scenario). One way of distinguishing the four scenarios in each experimental condition is by describing them as following one of four general storylines (i.e., plots), which differ in terms of how the task and student who solves it are selected, as well as how the correctness of the students' solution is discussed. The scenarios also differ in terms of the figure in the task – a feature that we use to title the storylines: the similar-triangles storyline, the trapezoid storyline, the isosceles-triangle storyline and the parallelogram storyline.

Each storyboard consists of 12 frames. During the first three frames, the class selects a GCA task to work on and the teacher either asks for a volunteer or selects a student to solve it, at the board. This occurs in one of four ways (depending on which storyline the given scenario follows), all of which were conjectured to be normative (e.g., the teacher in the scenario may accept a student's request to review a problem from the homework given the day before or may choose a problem that they conjecture might challenge the students; they may request that a particular student share their solution or ask for a volunteer). In the following three frames of each storyboard, the teacher puts the problem on the board and asks for the selected student to present their solution. In the CFCT and CFBT scenarios, the task complies with the GCAF norm, while in the BFBT scenarios, it breaches the GCAF norm by involving algebraic expressions that imply that the length of one of the sides of the figure is less than or equal to zero (and, therefore, that the figure does not exist). In the following three frames, the selected student writes a correct equation on the board, after which the teacher asks the student what theorem or property they

used to set-up that equation. In all cases, the student identifies a correct theorem but, whereas in the CFCT and CFBT scenarios the teacher then affirms the student, in the BFBT scenarios the teacher breaches the GCAT norm by saying something that conveys that they expected the student to write the theorem or property on the board (e.g., the teacher says “Why was that not written on the board? Please always write down the properties you use to justify your work.”). In the last three frames, the student finishes correctly solving the problem and asks the teacher to help them determine whether their solution is correct. This occurs in one of two ways, both of which were conjectured to be normative - the teacher either asks the class whether they think the solution is correct (which occurs in two of the storylines) or asks a specific student the same question (which occurs in the other two storylines). In the CFBT and BTBF scenarios, the theorem or property used to set up the first equation is also written on the board, in response to the teacher’s request.¹

For each item set, and hence for each scenario, the questionnaire contained seven open-response and five closed-response questions. After being shown the first six frames of a storyboard, participants were asked, “what did you see happening in this first segment of the scenario?” and provided with an open box in which they could type their response. Asking this type of question (a prompt for participants to describe what they notice about a given scenario) permits one to observe what participants tacitly expect will occur in those situations. In line with the notion of breaching experiments, we would expect that most participants who saw a scenario where GCAF had been breached would remark that breach. After that open-ended question, participants were asked to rate the appropriateness of the teacher's actions in the first three frames of the storyboard, using a 6-point Likert scale, and asked to explain their rating in an open-response field. The same appropriateness questions were asked about the second group of three frames of the storyboard. From the explanations of their ratings, we also expected to see evidence of participants’ recognition of the breaches of the GCAF norm as well as to learn why some teachers might disagree with certain breaches, while others might deem them justifiable.

Participants then saw the second half of the scenario (the third and fourth segments) and were asked the same three open-response and two closed-response questions about those segments. Last, they were asked to rate the appropriateness of the teacher's facilitation of the work on the problem throughout the scenario (again, using a 6-point Likert scale) and to explain their answer (in an open response field). This last question was posed, in particular, to provide participants with an opportunity to remark on breaches of the GCAF norm, in the chance that they had not realized that the task was not normative (if it was not) when it was first put on the board, but realized it once the student finished solving it.

Each participant was randomly assigned to one of three groups, each of which was assigned four item sets (two of one condition and two of another), as follows:

- Group 1 was assigned two CFCT item sets, one that followed the similar-triangles storyline and one that followed the trapezoid storyline, as well as two CFBT item sets, one that followed the isosceles-triangle storyline and one that followed the parallelogram storyline.
- Group 2 was assigned two CFBT item sets, one that followed the similar-triangles storyline and one that followed the trapezoid storyline, as well as two BFBT item sets,

¹ A table describing the four storylines in more detail is included in an extended version of this paper, available on Deep Blue.

one that followed the isosceles-triangle storyline and one that followed the parallelogram storyline.

- Group 3 was assigned two BFBT item sets, one that followed the similar-triangles storyline and one that followed the trapezoid storyline, as well as two CTCF item sets, one that followed the isosceles-triangle storyline and one that followed the parallelogram storyline.

Data analysis

As each participant was assigned two item sets of the same condition, we used mixed effects regression models (Agresti & Finlay, 2009) to analyze the closed-response data, using MemberID (a variable used to keep track of which responses were associated with which participant) as the random effect. The outcome variable in each model was the set of responses to one of the five Likert-scale items that asked participants to rate the appropriateness of the teacher's action (in each of the four segments of the scenario, then overall). To be able to compare ratings of scenarios of different experimental conditions and scenarios of different storylines, we created two variables - Condition (with values CFCT, CFBT, BFBT) and Storyline (with values similar-triangles, trapezoid, isosceles-triangle, parallelogram) – applied these to code each of the closed-responses and, after dichotomizing each, used those dichotomous variables as dependent variables in each model. The CFCT condition was used as the reference group for the CFBT and BFBT conditions, because the CFCT scenarios were designed as control scenarios (i.e., neither norm was breached in them). The choice to have the similar-triangles storyline as the reference group was arbitrary, as all scenarios were designed to be normative except for whether or not they complied with either the GCAF or GCAT norms.

We hypothesized that, when controlling for the Condition variable, there would be no significant differences between the mean ratings of the scenarios of different storylines, in any of the models, as we conjectured that the teachers' actions in each were equally appropriate (as they were designed to be normative, outside of the moments when each of the norms were at issue). Similarly, we also conjectured that there would be no significant differences between the mean rating of segments 1 and 4 of the scenarios, when controlling for the Storyline variable. The only significant differences we expected to observe were in the mean ratings of segments 2 and 3 of the scenarios, because they each relate to part of the text where one of the norms was breached in some items, but not in others. Specifically, since the GCAF norm was only breached in segment 2 of the BFBT scenarios, we expected that it would be rated significantly lower than segment 2 in CFCT (the reference group) scenarios, on average. As the GCAT norm was breached in segment 3 of both the CFBT and BFBT scenarios, we expected that it would be rated significantly lower than segment 3 in the CFCT scenarios, on average.

Last, to evaluate our hypothesis that teachers would recognize the breaches of our two hypothesized norms, we created two dichotomous codes – one representing that there was evidence of recognition (versus non-recognition) of the breach of the GCAF norm and one representing that there was evidence of recognition (versus non-recognition) of the breach of the GCAT norm – and applied each to all open-response items. The following is an example of a response coded for recognition of a breach of the GCAT norm: “Kids solves it but doesn't write justification. Teacher tells kid (with different word choice) to write justifications. Kid does it and we move on.”

Each participants was then given two scores – one indicating whether there was evidence of recognition of a breach the GCAF norm in any of their open responses and another indicating whether there was evidence of recognition of a breach of the GCAF norm in any of their open

responses. A series of z-tests were then conducted to determine whether most participants who were assigned each item that contained a breach of one or both norms recognized those breaches.

Results

Results related to research question 1

In terms of the results of the twelve z-tests of the proportion of participants who recognized the breach of the GCAT norm in each open-response item (against the null hypothesis of 50% recognition), as predicted, no participants recognized a breach in any of the CFCT-condition items, as the norm was not breached in those scenarios. The proportion of participants who recognized the breach in each of the CFBT-condition and BFBT-condition items, except one, ranged from 0.45 to 0.73 (but none of those values were statistically significant at the level of 0.05).

Further, when coding the open-response data for recognition of the breaches of the GCAT norm, we noticed that there were several participants who noted that the teacher requested that the student state the theorem or property that they used to set-up the equation. Consequently, we also conducted a z-test of the proportion of participants that recognized this request in relation to each item. The results indicate more than 50% recognition in all the CFBT-condition and BFBT-condition items (in some cases, significantly more). In fact, even the proportion of participants that recognized this request in each of the four CFCT-condition items was 0.82, 0.43, 0.47 and 0.33.

One surprising result was that, although all BFBT scenarios contained a breach of the GCAF norm, there was only evidence that one participant recognized one of those breaches. Further, this evidence was in their response to the rating of the third segment of the scenario, when they wrote: “the teacher is allowing the student to discover that there is no solution to the problem”. Therefore, there was no recognition of the breach of the GCAF norm when the problem was being written on the board (despite this being done over three frames). As we discuss in the next section, this result suggests the need for deeper exploration of the GCAF norm.

Results related to research questions 2 and 3

The results of the five mixed-effect regression analyses (one for each of the ratings of the four segments of the scenario and one for the rating of the teacher’s facilitation of the work on the problem throughout the scenario) are summarized in table 1 and generally confirm most of our hypotheses.

As indicated in table 1, when controlling for the Condition variable, we see that all but two of the coefficients are not significant. This supports our earlier claim that, outside of the breach or compliance with one or both of the norms, participants rated the teachers’ actions in the scenarios as being similarly appropriate. On the other hand, the two significant Storyline coefficients suggest that this might not be the case. We will discuss this point further in the next section.

When controlling for the Storyline variable, the coefficients for the Segment-3 ratings and the Overall ratings of the CFBT scenarios and BFBT scenarios are negative and significant. This indicates, as hypothesized, that ratings of the segment of the scenarios in which the GCAT norm was breached would be rated lower, on average, than the equivalent segments of scenarios in which that norm was not breached (CFCT scenarios), and that the same would consequently be true for the overall rating of the scenarios. However, neither the rating of segment 3 or the overall rating associated with BFBT scenarios was significantly different than those of the CFBT scenarios. Although this does not provide us with evidence to believe that participants’ reactions

to breaches of the GCAT norm are influenced by the reactions to breaches of the GCAF norm, as we also discuss in the next section, lack of recognition of the GCAF norm could be the explanation for this.

Table 1: Summary of Mixed Effect Regression Analyses for Variables Predicting Ratings of Each Segment of the Scenarios and the Overall Rating

	<u>Seg-1 rating</u> B(SE)	<u>Seg-2 rating</u> B(SE)	<u>Seg-3 rating</u> B(SE)	<u>Seg-4 rating</u> B(SE)	<u>Overall rating</u> B(SE)
Fixed effects					
Trap storyline	-0.55(0.28)*	0.14(0.21)	0.02(0.18)	0.06(0.24)	0.11(0.16)
Iso-tri storyline	-0.22(0.24)	0.04(0.25)	0.03(0.23)	-0.26(0.21)	0.27(0.17)
Parall storyline	-0.04(0.23)	0.27(0.21)	-0.02(0.27)	-0.03(0.22)	0.45(0.22) *
CFBT cond	0.46(0.24)*	0.60(0.20)**	-0.95(0.23)***	-0.11(0.18)	-0.63(0.24)**
BFBT cond	0.21(0.22)	0.14(0.20)	-0.98(0.29)***	-0.23(0.18)	-0.56(0.20)**
Constant	4.19(0.23)***	3.72(0.21)***	5.00(0.20)***	4.78(0.21)***	4.58(0.22)***
Random effects					
Constant	0.41(0.23)	-0.19(0.14)	0.24(0.16)	0.32(0.13)*	0.23(0.15)
Residual	0.08(0.09)	0.10(0.07)	0.02(0.09)	0.01(0.10)	0.21(0.08)
N	158	158	158	158	158

Standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Also in line with our hypotheses, the mean ratings of segment 4 in all scenarios were similar, which was expected, as no norm was breached in that segment of any of the scenarios. In contrast, however, segment 1 and segment 2 of the CFBT scenarios were rated significantly higher, on average, when controlling for the Storyline variable, than those segments in the CFCT scenarios, which contrasted with our hypothesis that they would be rated similarly (as both of those segments in all of those scenarios were also designed to represent normative instruction).

Discussion: Potential Revisions to the Instrument

Although the results of the study are mixed, in the sense that the instrument provided us with evidence that the GCAT norm is, in fact, a norm of the instructional situation of GCA, but did not allow us to conclude the same about the GCAF norm, nor to detect any relationship between the way that participants reacted to the two norms, the results do provide us with ideas for future research. For one, we would argue that the general design of the instrument is promising. The “what did you see happening...” questions provided us with evidence that one of the hypothesized norms exists. Another affordance of the open-response questions was that they not only allowed us to collect evidence in support of our hypothesized norms, but also allowed us to consider whether and how to revise these hypotheses. Although the proportion of participants that recognized breaches of the GCAT norm in some of the scenarios was as high as we expected, there were scenarios for which the proportion was lower. Of course, this could be a consequence of the small sample size or the representativeness of the sample, but the proportions of participants that recognized the teachers request that the student state the theorem or property at least suggests that the norms might have been slightly different than we first conjectured. For example, it could be that the norm is in fact that students are not expected to write or state the theorem or property that they used to set-up their equation(s). This alternative is supported by the fact that many of the participants’ attention were drawn to the request to state the theorem of property, even in the CFCT scenarios.

In terms of the lack of recognition of the GCAF norm, although it could indicate that the norm does not exist, we argue that this is more likely a consequence of at least one of the following two issues with the scenarios. The first is that, in order to detect that the norm was breached, a participant would have likely had to work through the problem, which they might not ordinarily have to, if they assumed that the task was normative, as it had been put on the board by the teacher in the scenarios. Alternatively, there is also evidence in the open-response data that there were more distracting aspects of the scenarios than whether or not the norm was breached. In particular, many participants commented that the teacher took too long (3 frames) to write the problem on the board and that they should have instead used a document camera or have the student write the problem on the board, while the teacher circulated. We found the suggestion of having the student put the problem on the board to be especially helpful, as we conjecture that the teacher would also more likely analyze a problem if the student was the one putting it on the board, and are considering revising the items to integrate this change.

Last, as we are more likely to detect a relationship between the two norms if participants recognize the breaches of the first norm, we are considering adding a question, after they evaluate the first half of the scenario but before they evaluate the second half, that will ask participants to rate the appropriateness of the task in the scenario, expecting that this will also require them to analyze the GCA task (if they did not do so when it was put on the board). Similar to our other rating questions, their rating of the task and their explanation of that rating could also provide us with some evidence that teachers recognize the GCAF norm, even if breaches are not remarked in the first three open-responses.

Acknowledgments

The research reported in this article had the support of the National Science Foundation, through grant DRL-0918425 to P. Herbst. All opinions are those of the authors and don't necessarily reflect the views of the Foundation.

References

- Agresti, A., & Finlay, B. (2009). *Statistical methods for the social sciences, 4th edition*. Upper Saddle River, New Jersey: Pearson.
- Dimmel, J. K. (2015). *How do teachers expect students to represent mathematical work? A study of teachers' recognition of routine ways that proofs are presented and checked in high school geometry* (Unpublished doctoral dissertation). The University of Michigan, Ann Arbor, MI.
- Garfinkel, H. (1963). A conception of and experiments with "trust" as a condition of concerted stable actions. In O.J. Harvey (ed.), *Motivation of Social Interaction* (381-392). New York: The Ronald Press Company.
- Herbst, P. (2006). Teaching geometry with problems: Negotiating instructional situations and mathematical tasks. *Journal for Research in Mathematics Education*, 37, 313-347.
- Herbst, P., & Chazan, D. (2012). On the instructional triangle and sources of justification for actions in mathematics teaching. *ZDM*, 44(5), 601-612.
- Herbst, P., Kosko, K., & Dimmel, J. (2013). *How are geometric proof problems presented? Conceptualizing and measuring teachers' recognition of the diagrammatic register*. Proceedings of the 2013 Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education. Retrieved from Deep Blue at The University of Michigan: <http://hdl.handle.net/2027.42/97761>.
- Hsu, H. (2010). *The study of Taiwanese students' experiences with geometric calculation with number (GCN) and their performance on GCN and geometric proof (GP)* (Unpublished doctoral dissertation). The University of Michigan, Ann Arbor, MI.
- National Council of Teachers of Mathematics (NCTM). (2014). *Principles to Actions: Ensuring Mathematical Success for All*. Reston, VA: Author.
- Schoenfeld, A. H. (2010). *How We Think: A Theory of Goal-Oriented Decision Making and its Educational Applications*. New York: Routledge.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Wadsworth Cengage learning.