PARCC

Operational Study 4: Accessibility of New Items/Functionality Component 3 Report

Jeffrey Steedle and Amy LaSalle
Pearson
October 31, 2016

Introduction

PARCC Operational Study 4 Component 3 was designed to compare performance on PARCC mathematics field-test items for grade 3 taken with and without a drawing tool. For the 2016 testing window, five field-test items were selected to have the directions edited to allow students to provide a drawing as part of responses and directions on how to include a drawing. For those students, using the drawing tool was optional, and a drawing was not required for earning full credit. The standard variant of these items did not include the directions or the tool option. Both variants were included in the embedded online field-test forms to determine if there was a significant difference in performance. If no significant differences were found, then there would be evidence to allow the inclusion of the drawing response interaction on existing items. Because the study items were located in field test positions, scores on these items did not contribute to overall assessment scores, so no students were advantaged or disadvantaged by access to the drawing tool.

All mathematics items available to be embedded into the Spring 2016 forms were reviewed and approved by PARCC. No new items were created for this study, only variant items with modified direction lines. This report summarizes the quantitative and qualitative evidence bearing on the comparability of the study items administered with and without a drawing tool.

This report includes the following sections: Study Items to describe the items administered for this study; Sample Representation to describe how the groups of students taking the study items with and without access to the drawing tool were similar or dissimilar in terms of demographic variables; Scoring Methodology to explain the special steps taken when scoring responses for students who used the drawing tool; Scorer Reliability to provide evidence that scorers had consistent evaluations of responses to the study items; Drawing Tool Usage to summarize how frequently the drawing tool was used and whether the drawings affected scoring; Performance Comparisons to summarize how average scores differed between students with and without access to the drawing tool; Differential Item Functioning to indicate whether there was statistical evidence of potential item bias on the study items; Item-Total Correlations to show whether the study items were more or less discriminating when students had access to the drawing tool; Time-on-Task to compare the length of time spent on the study items with and without access to the drawing tool; and Observations to summarize comments made by scoring directors and content specialists. This report concludes with recommendations based on results and next steps for future research and implementation.

Study Items

For this study, the instructions for five grade 3 mathematics items were modified to inform students that they had the option to use a drawing as part of their responses and how to use the drawing tool. Some of the study items were multi-part items, with only one part allowing use of the drawing tool.

- M02861 is a single-part item associated with evidence statement 3.C.6-2 scored on a 0–3 point rubric. M02861D is the drawing tool version of this item.

- 4177-M03909 is a multi-part item associated with evidence statement 3.C.4-4 worth a total of 3 points. The hand-scored part was scored on a 0–2 point rubric. 4177-M03909D is the drawing tool version of this item.
- M300083P is a multi-part item associated with evidence statement 3.C.2 worth a total of 4 points. The hand-scored part was scored on a 0–3 point rubric. M300083PD is the drawing tool version of this item.
- 4220-M04135 is a multi-part item associated with evidence statement 3.D.1 worth a total of 3 points. The hand-scored part was scored on a 0–2 point rubric. 4220-M04135D is the drawing tool version of this item.
- M04129 is a single-part item associated with evidence statement 3.D.1 scored on a 0–3 point rubric. M04129D is the drawing tool version of this item.

Sample Representation

Given that assessment forms were randomly assigned (or "spiraled"), the students taking the study items with and without drawing tools should be randomly equivalent. That is, their characteristics, including distributions of demographic variables and mathematics ability, should be nearly the same. Table 1 shows the proportion of students in each study condition (No Tool or Drawing Tool) in demographic groups. Cohen's *h* was calculated as an effect size (i.e., difference in standard deviation units). In all cases, the magnitude of *h* was below 0.09 (average magnitude of 0.02), suggesting overall similarity between the groups. As a consequence of this similarity, observed differences in performance on the study items can be attributed to the drawing tool. A demographic comparison for each item is provided in the Appendix.

Table 1
Demographic Comparison for All Items

	Proportion of Sample			
	No Tool	Drawing Tool	Diff.	Cohen's *h*
Female	0.498	0.502	0.003	0.007
Male	0.499	0.501	0.002	0.004
American Indian	0.508	0.492	-0.016	-0.033
Asian	0.503	0.497	-0.005	-0.010
Black	0.495	0.505	0.011	0.022
Hispanic	0.494	0.506	0.012	0.023
Pacific Islander	0.482	0.518	0.036	0.073
White	0.501	0.499	-0.002	-0.005
Multi-Race	0.490	0.510	0.020	0.040
Not Provide	0.505	0.495	-0.010	-0.020
ELL Unknown	0.505	0.495	-0.009	-0.019
Not ELL	0.499	0.501	0.001	0.003
ELL	0.491	0.509	0.018	0.035

Econ. Dis. Unknown	0.499	0.501	0.001	0.002
No Econ. Dis.	0.499	0.501	0.002	0.003
Econ. Dis.	0.479	0.521	0.042	0.083
Disability Unknown	0.498	0.502	0.004	0.007
Disability (504)	0.502	0.498	-0.004	-0.007
Disability (IEP)	0.511	0.489	-0.022	-0.045
No Disability	0.498	0.502	0.005	0.010

Scoring Methodology

Pearson's scoring team established methodology and criteria with PARCC for scoring the standard variant and tool variant responses. The method is described below.

1. During rangefinding, representative student responses for both the standard-variant and tool-variant responses were selected to represent the range of responses at each score point. Training paper sets were created for each response type.
2. Nine scorers participated in the scoring of the drawing tool study items and the collection of information. Table 2 below shows the distribution of scorers who scored each item.

Table 2
Scorer Participation by Item

Item ID	*No of Scorers*
M03_4177_M03909	8
M03_4177_M03909D	7
M03_4220_M04135	6
M03_4220_M04135D	7
M03_M02861	7
M03_M02861D	7
M03_M04129	9
M03_M04129D	8
M03_M300083P	9
M03_M300083PD	9

3. Scorers were initially trained using the student responses for the standard variant items. Validity and inter-rater reliability (IRR) metrics were collected on each standard-variant item to validate the correct and consistent application of the rubric to the range of student non-variant responses.
4. Scorers were trained with the tool-variant student responses. Validity and IRR metrics were collected on the tool-variant responses to validate the correct and consistent application of the rubric to the range of student tool-variant responses.
5. For each response in which students used the tool, the scorers captured the following information:

a. Score point
b. Scoring impact – the extent to which the drawing had a positive, negative, or neutral effect on the student score.
    - **Positive Effect**. The response in the drawing box positively contributed to the student score
    - **Negative Effect.** The response in the drawing box negatively impacted or lowered the student score. For example, the response in the drawing box contained an error that would not have been present if the drawing were not included.
    - **Neutral Effect.** The response in the drawing box did not increase or decrease the achieved score.
c. Qualitative Observations about the nature and content of the drawings, whether the work reinforced or clarified the math responses for Part A, whether the work introduced errors not present in Part A, and observations about the students' ability to use the tool to respond to the question.
d. Summary of number of responses that attempted use of the drawing tool

Scorer Reliability

The valid interpretation of results from this study depends on the extent to which scorers applied scores to student responses that validly reflect the students demonstrated performance. Evidence of scorer reliability supports the validity of score interpretations because it indicates that scorers had accurate and similar perceptions about the quality of responses. For this study, it would be helpful to establish that scorer reliability was similar on the versions of the items with and without the drawing tool.

To continually monitor scoring quality, scorers are randomly seeded "validity" papers with scores established by expert scorers. When scorers are scoring accurately, their scores on the validity papers should be nearly identical to the established scores. Indeed, exact agreement on the validity papers exceeded 90% in all cases except M300083PD (with the drawing tool; Table 3). Even so, 84.9% exact agreement would be considered adequately high. The scorers never differed by more than 1 point from the established scores (Exact + Adjacent Validity %).

Inter-rater reliability is estimated by comparing scores from different operational scorers. As shown in Table 2, scorers agreed exactly on student scores at least 85.8% of the time. Exact plus adjacent inter-rater reliability was 100% or nearly so in every case. All observed levels of exact and adjacent inter-rater reliability would be considered adequately high to report scores.

Table 3
Inter-Rater Reliability

Item	Validity Read Count	Exact Validity %	Exact + Adj. Validity %	Reliability Read Count	Exact Agreement %	Exact + Adj Agreement %
M02861	74	97.3%	100.0%	329	89.7%	97.6%
M02861D	75	98.7%	100.0%	326	92.6%	100.0%

4177_M03909	76	96.0%	100.0%	162	85.8%	100.0%
4177_M03909D	74	91.8%	100.0%	163	88.3%	99.4%
M04129	76	96.1%	100.0%	332	89.2%	100.0%
M04129D	70	92.9%	100.0%	333	91.0%	100.0%
4220_M04135	71	95.8%	100.0%	326	95.1%	100.0%
4220_M04135D	69	97.1%	100.0%	326	96.9%	100.0%
M300083P	72	98.6%	100.0%	322	91.9%	100.0%
M300083PD	73	84.9%	100.0%	325	85.8%	98.2%

Drawing Tool Usage

Not all students with access to the drawing tool actually used it. For this study, scorers in Pearson's Performance Scoring Center kept a record of the number of students who used the drawing tool, and they made judgments about whether the students' drawings had positive or negative effects on their scores (Table 4). Depending on the item, between 37.0% and 64.1% of students used the drawing tool. The drawings had a perceived positive impact more often than a perceived negative impact, but scorers did not often perceive that drawing tool usage affected scores. When interpreting these results, one must consider that, even if a student used the drawing tool and the scorer perceived a positive impact, the student might have received a similar score without access to the drawing tool (e.g., by providing additional written explanation). It is notable that the greatest percentage of perceived positive impacts (12.5%) occurred for item 4177-M03909D, the item for which students with access to the drawing tool tended to perform better.

Table 4
Summary of Drawing Tool Usage

Item	Total	Used Drawing Tool	Tool Had Positive Effect	Tool Had Negative Effect
M02861D	1529	980	57	39
		64.1%	3.7%	2.6%
4177-M03909D	1499	941	187	16
		62.8%	12.5%	1.1%
M300083PD	1500	555	56	0
		37.0%	3.7%	0.0%
4200-404135D	1501	751	9	0
		50.0%	0.6%	0.0%
M04129D	1529	813	58	10
		53.2%	3.8%	0.7%
All	7558	4040	367	65
		53.5%	4.9%	0.9%

Performance Comparisons

Table 5 shows the distributions of scores in the groups of students with and without access to the drawing tool. If item difficulty was similar for students in the both groups, the percentage of students in each group at each score point would be similar. Overall, results suggest similar item difficulty for students with and without access to the drawing tool. The only notable deviation from this trend was item 4177-M03909, whose score distribution indicated better performance for students with access to the drawing tool. Note also that the rates of omitting were similar in the two groups.

Table 5
Score Distributions (Hand-Scored Parts Only)

Item	Group	Total	0	1	2	3	Omit
M02861	No Tool	1530	59.5%	22.8%	10.4%	4.9%	2.4%
	Drawing Tool	1529	59.7%	24.7%	9.5%	4.3%	1.7%
4177-M03909	No Tool	1501	73.8%	20.3%	2.3%	*	3.6%
	Drawing Tool	1499	68.9%	21.5%	6.4%	*	3.2%
M300083P	No Tool	1499	32.9%	17.1%	25.6%	22.0%	2.5%
	Drawing Tool	1500	33.1%	18.9%	23.2%	22.1%	2.7%
4220-M04135	No Tool	1500	86.1%	3.2%	6.5%	*	4.2%
	Drawing Tool	1501	85.4%	2.8%	6.6%	*	5.1%
M04129	No Tool	1530	43.8%	22.3%	24.9%	7.4%	1.6%
	Drawing Tool	1529	44.2%	21.4%	27.1%	5.9%	1.4%

* The hand-scored portions of Items 4177-M03909 and 4220-M04135 were worth 0–2 points, whereas the hand-scored portions of the other study items were worth 0–3 points.

The trend illustrated by Table 5 was confirmed with statistical testing. Specifically, a *t*-test was conducted to detect statistically significant differences in average performance between students with and without access to the drawing tool. Only the *t*-test for item 4177-M03909 indicated a statistically significant difference (Table 6). The corresponding effect size indicated that students with access to the drawing tool had an average item score that was 0.17 standard deviations higher than students with no drawing tool.

Table 6
Item Performance Comparison (Hand-Scored Part Only)

	Average Item Score					
Item	No Tool	Drawing Tool	Difference	*t*	*p*	Effect Size
M02861	0.598	0.577	-0.021	-0.670	.503	-0.025
4177-M03909	0.259	0.354	0.095	4.676	.000	0.174
M300083P	1.377	1.353	-0.024	-0.553	.580	-0.021
4220-M04135	0.170	0.169	-0.001	-0.052	.959	-0.002

M04129	0.958	0.946	-0.013	-0.348	.728	-0.013

Differential Item Functioning

Differential item functioning (DIF) indicates whether certain groups of students (e.g., ethnic groups, English language learners) performed unexpectedly well or poorly on an item. For example, if English language learners performed worse on an item than expected (given their performance on the rest of the test), that would be indicated by a DIF analysis. If the drawing tool was to be implemented operationally, it should be demonstrated that the drawing tool does not impart an unfair advantage to certain groups of students. Ten DIF analyses were conducted for each study item (male/female, White/Black, White/Hispanic, White/Asian, White/Native American, White/Pacific Islander, White/Multi-Race, Non-ELL/ELL, students with no disabilities/students with disabilities, and no economic disadvantage/economic disadvantage. None of the DIF analyses indicated evidence of potential item bias.

Item-Total Correlations

Generally, an item-total correlation (e.g., polyserial correlation) indicates the degree to which an item differentiates between students of lower and higher ability, which is a desirable trait of test items. If the use of the drawing tool minimizes a construct-irrelevant barrier to demonstrating proficiency (as intended), a higher item-total correlation may be observed for students with access to the drawing tool. The polyserial correlations for the two groups were compared using the Fisher *z*-test to evaluate statistical significance and Cohen's *q* to evaluate practical significance (Table 7). For the study items, the polyserial correlation ranged from 0.667 to 0.861, which indicates that the study items discriminated well between students of lower and higher ability. The statistical evidence suggests that the polyserial correlations were similar for students with and without access to the drawing tool. None of the differences in correlations would be considered statistically significant ($z < -1.96$ or $z > +1.96$), and all of the effect sizes (q) fell below 0.03, which would be considered negligible.

Table 7
Item-Total Correlations

	Polyserial Correlation			
Item	No Tool	Drawing Tool	*z*	*q*
M02861	0.861	0.853	-0.362	-0.013
4177-M03909	0.684	0.667	-0.372	-0.014
M300083P	0.744	0.725	-0.490	-0.018
4220-M04135	0.705	0.737	0.793	0.029
M04129	0.715	0.727	0.300	0.011

Time-on-Task

Table 8 lists the average time spent on the study items with and without access to the drawing tool. On average, students spent approximately 2.5 to 3.0 minutes longer on the study items when they were given access to the drawing tool.

Table 8
Time-on-Task Comparison

	Time-on-Task (Seconds)		
Item	No Tool	Drawing Tool	Difference
M02861	458	625	167
4177-M03909	543	724	181
M300083P	497	645	148
4220-M04135	493	641	148
M04129	480	656	176

Since students tended to spend longer on the study items when a drawing tool was available, those students would have had less time available for the remainder of the assessment. Conceivably, those students could have performed worse on the overall assessment due to time pressures. To investigate this possibility, mean performance on other items for students with and without access to the drawing tool were compared. This analysis was conducted with operational items in the same unit as the study items because these units included the same items (other than the study items) for students with and without access to the drawing tool. Operational items in other units were not included because they sometimes differed, and that would make the comparison difficult to interpret. As shown in Table 9, only the mean difference in unit scores for students who took item M300083P was statistically significant ($p < .01$). Considering the effect size, this means that students with access to the drawing tool scored an average of 0.0423 standard deviations lower on the operational items (from the same unit as the study item) than students with no access to the drawing tool. All effect sizes in Table 9 would be considered very small or negligible, but it is notable that average unit scores were consistently lower for students with access to the drawing tool (i.e., the differences and their corresponding *t* statistics and effect sizes were always negative).

Table 9
Unit Performance Comparison

	Average Unit Score					
Item	No Tool	Drawing Tool	Difference	*t*	*p*	Effect Size
M02861	6.28	6.23	-0.05	-1.440	.150	-0.020
4177-M03909	5.52	5.46	-0.06	-1.397	.163	-0.019
M300083P	5.59	5.46	-0.13	-3.091**	.002	-0.042
4220-M04135	5.46	5.43	-0.03	-0.729	.466	-0.010
M04129	6.21	6.17	-0.04	-1.105	.269	-0.015

** $p < .01$

Observations

The study scoring and observations were conducted by Pearson's most experienced scoring staff, including the scoring directors and supervisors who oversee the scoring process. The observations of this staff are recorded in Table 10. These comments provide valuable qualitative description of the use and usability of the drawing tool.

Table 10
Scoring Director & Supervisor Comments

M02861D	The drawing box helped some of the students explain how to use the number line to find the number of minutes in three sets. Some students did use the drawing box to do the other work for the prompt as well, doing equations and standard algorithms to find the number of minutes in a set or in finding the finish time for the basketball players. I would recommend the use of the drawing box for this item.
4177-M03909D	The drawing box seemed to help the students explain how to correct or create new models in order to obtain the second score point of this item. For this item, I think the drawing box was appropriate. Students referenced that they had colors available to them in the drawing box. We were only able to see differing shades of grey when they used the available colors. One thing that might have been a help for the students might have been an option to make geometric shapes as a button on the control panel.
M300083PD	The drawing box seemingly made some students feel like they have to include a drawing. Therefore it became an exercise in that they tried to make it match their answer and did not use it to explain or get the answer. Since an equation was not required for this item, some students were able to use the drawing box to count and to show their work. Calling it a work space instead of drawing tool might be helpful to encourage students to continue or start their work there. Changing the order and putting the drawing box first or side by side might help students see it as something they can use to help solve the problem. It would be helpful for the drawing box to be bigger. When students try to write numbers with the pencil tool, they tend to run out of space.

4200-404135D	No noticeable difference in how the students answered the question in this item as opposed to how they answered the regular item. Some students did use the drawing box to do the work (both equations and standard algorithms). A large number of drawings that were of children playing instruments, sometimes with valid responses in the text box and other times not. I would not recommend the use of the drawing box for this item.
M04129D	For some students the drawing tool helped and they received an extra point while others attempted work that did not change their score. Some students repeated the answer provided in the answer box. Many students drew pictures of flowers, wasting valuable test time. Maybe calling it a “work space” would prevent students from feeling like they needed to draw a picture. In this item, the drawing box could have been used by students to set up a vertical algorithm to borrow correctly on the subtraction problem, 70 – 56 =14, but few students used it that way. Those that showed how to divide the flowers into the vases would not receive credit because an equation is required in the prompt. It was very rare for a student to lose a point due to information provided in the drawing box. This occurred when the student provided work that contradicted work that was provided in the answer box, thereby providing correct and incorrect work. It would be helpful for the drawing box to be bigger. When students try to write numbers with the pencil tool, they tend to run out of space.

Scoring directors and supervisors provided additional general commentary:

- Responses where the drawing did not help nor hurt the student were considered neutral. The student did spend time placing some drawing or work in the drawing box, but the final score was not increased or decreased due to the drawing provided. The majority of work in the drawing boxes across all prompts was neutral.
- Many students drew "pictures" in the drawing box. Relabeling of the box might help decrease the number of irrelevant pictures.
- Use of the pencil tool is challenging and difficult to make small numbers, or even lines. Increasing the size of the box might help, or providing line segments and other geometric figures that students can place in the box without use of the pencil tool might help.

Recommendation

According to plans for Study 4 Component 3, if the results of this study suggested a neutral-positive impact of the drawing tool, the recommended course of action would be to (1) allow the study items to be administered operationally with a drawing tool, (2) identify additional grade 3 items to field-test with a drawing tool, and (3) conduct a similar statistical investigation of the drawing tool using field-test items at grades 4 and 5. In this study, 4 items suggested the drawing tool had no impact on student performance, and 1 item suggested a positive impact. Thus, the preliminary recommendation of this study is to continue pursuing the use of the drawing tool.

Next Steps

In August, Pearson presented the preliminary results of Study 4, Component 3 to the State Leads. The State Leads voted to move forward with the preliminary recommendations of the study to make the grade 3 drawing tool items eligible for placement on tests in spring 2017, to field test additional drawing tool items at grade 3, and to extend the study to determine if results were generalizable at grades 4 and 5. Based on that approval, an additional four drawing tool items will be field tested at grade 3; and six items (three field test items each at grades 4 and 5) will go forward as Component 4 of the PARCC Study 4 in spring 2017. The preliminary results of that study will be available in August 2017.

Appendix

Table A1

Demographic Comparison for Item M02861

	Proportion of Sample			
	No Tool	Drawing Tool	Diff.	Cohen's *h*
Female	0.492	0.508	0.015	0.031
Male	0.504	0.496	-0.009	-0.017
American Indian	0.500	0.500	0.000	0.000
Asian	0.498	0.502	0.004	0.008
Black	0.491	0.509	0.018	0.036
Hispanic	0.495	0.505	0.009	0.019
Pacific Islander	0.500	0.500	0.000	0.000
White	0.501	0.499	-0.003	-0.005
Multi-Race	0.494	0.506	0.011	0.023
Not Provide	0.506	0.494	-0.011	-0.023
ELL Unknown	0.500	0.500	0.000	0.000
Not ELL	0.500	0.500	0.001	0.001
ELL	0.484	0.516	0.031	0.062
Econ. Dis. Unknown	0.501	0.499	-0.001	-0.002
No Econ. Dis.	0.499	0.501	0.003	0.005
Econ. Dis.	0.462	0.538	0.076	0.153
Disability Unknown	0.497	0.503	0.007	0.014
Disability (504)	0.515	0.485	-0.029	-0.058
Disability (IEP)	0.520	0.480	-0.040	-0.081
No Disability	0.496	0.504	0.007	0.015

Table A2

Demographic Comparison for Item 4177-M03909

	Proportion of Sample			
	No Tool	Drawing Tool	Diff.	Cohen's *h*
Female	0.502	0.498	-0.004	-0.008
Male	0.492	0.508	0.016	0.032
American Indian	0.548	0.452	-0.096	-0.193
Asian	0.514	0.486	-0.027	-0.055
Black	0.503	0.497	-0.006	-0.013
Hispanic	0.492	0.508	0.016	0.033
Pacific Islander	0.486	0.514	0.029	0.057
White	0.495	0.505	0.010	0.020
Multi-Race	0.488	0.512	0.025	0.049
Not Provide	0.498	0.502	0.004	0.007
ELL Unknown	0.540	0.460	-0.079	-0.159
Not ELL	0.495	0.505	0.010	0.020
ELL	0.505	0.495	-0.010	-0.021
Econ. Dis. Unknown	0.497	0.503	0.006	0.012
No Econ. Dis.	0.497	0.503	0.006	0.011
Econ. Dis.	0.491	0.509	0.017	0.035
Disability Unknown	0.496	0.504	0.008	0.016
Disability (504)	0.513	0.487	-0.027	-0.053
Disability (IEP)	0.506	0.494	-0.012	-0.024
No Disability	0.496	0.504	0.009	0.017

Table A3

Demographic Comparison for Item M300083P

	Proportion of Sample			
	No Tool	Drawing Tool	Diff.	Cohen's *h*
Female	0.500	0.500	0.000	0.001
Male	0.500	0.500	0.000	0.000
American Indian	0.475	0.525	0.051	0.102
Asian	0.490	0.510	0.019	0.039
Black	0.498	0.502	0.005	0.009
Hispanic	0.495	0.505	0.009	0.018
Pacific Islander	0.375	0.625	0.250	0.505
White	0.504	0.496	-0.008	-0.016
Multi-Race	0.500	0.500	0.000	0.000
Not Provide	0.508	0.492	-0.016	-0.033
ELL Unknown	0.503	0.497	-0.006	-0.013
Not ELL	0.501	0.499	-0.002	-0.004
ELL	0.486	0.514	0.027	0.055
Econ. Dis. Unknown	0.500	0.500	0.000	-0.001
No Econ. Dis.	0.501	0.499	-0.003	-0.006
Econ. Dis.	0.469	0.531	0.063	0.125
Disability Unknown	0.499	0.501	0.002	0.004
Disability (504)	0.496	0.504	0.007	0.015
Disability (IEP)	0.486	0.514	0.027	0.054
No Disability	0.502	0.498	-0.004	-0.008

Table A4

Demographic Comparison for Item 4220-M04135

	Proportion of Sample			
	No Tool	Drawing Tool	Diff.	Cohen's *h*
Female	0.495	0.505	0.009	0.018
Male	0.501	0.499	-0.003	-0.005
American Indian	0.514	0.486	-0.028	-0.057
Asian	0.504	0.496	-0.009	-0.017
Black	0.486	0.514	0.027	0.055
Hispanic	0.494	0.506	0.013	0.025
Pacific Islander	0.500	0.500	0.000	0.000
White	0.504	0.496	-0.009	-0.017
Multi-Race	0.458	0.542	0.084	0.169
Not Provide	0.499	0.501	0.002	0.004
ELL Unknown	0.491	0.509	0.017	0.034
Not ELL	0.499	0.501	0.002	0.004
ELL	0.493	0.507	0.014	0.028
Econ. Dis. Unknown	0.499	0.501	0.002	0.004
No Econ. Dis.	0.498	0.502	0.005	0.009
Econ. Dis.	0.502	0.498	-0.005	-0.010
Disability Unknown	0.501	0.499	-0.001	-0.003
Disability (504)	0.485	0.515	0.031	0.061
Disability (IEP)	0.537	0.463	-0.075	-0.150
No Disability	0.493	0.507	0.013	0.027

Table A5

Demographic Comparison for Item M04129

	Proportion of Sample			
	No Tool	Drawing Tool	Diff.	Cohen's *h*
Female	0.502	0.498	-0.005	-0.009
Male	0.498	0.502	0.004	0.009
American Indian	0.502	0.498	-0.005	-0.010
Asian	0.507	0.493	-0.014	-0.028
Black	0.495	0.505	0.010	0.020
Hispanic	0.495	0.505	0.011	0.021
Pacific Islander	0.550	0.450	-0.100	-0.200
White	0.501	0.499	-0.002	-0.005
Multi-Race	0.509	0.491	-0.018	-0.036
Not Provide	0.512	0.488	-0.025	-0.050
ELL Unknown	0.489	0.511	0.022	0.044
Not ELL	0.502	0.498	-0.003	-0.006
ELL	0.487	0.513	0.026	0.053
Econ. Dis. Unknown	0.501	0.499	-0.001	-0.003
No Econ. Dis.	0.501	0.499	-0.002	-0.004
Econ. Dis.	0.473	0.527	0.054	0.109
Disability Unknown	0.499	0.501	0.002	0.005
Disability (504)	0.500	0.500	0.000	0.000
Disability (IEP)	0.505	0.495	-0.011	-0.022
No Disability	0.500	0.500	0.000	-0.001