

Technical Summary of Preliminary Analyses of NAEP 2017 Writing Assessments

Reflecting the increasing use of digital devices in writing composition, the new NAEP Writing Framework developed for the 2011 assessment specifies that students' writing skills be assessed "...using word processing software with commonly available tools (NAEP Writing Framework, p. vi)." To implement this framework, NCES first administered digitally based writing assessments to students in grades 8 and 12 in 2011 and subsequently in 2017 to students in grades 4 and 8. Preliminary analyses of students' writing performance in the 2017 NAEP writing assessments revealed patterns that require further analysis. This document summarizes the results of these preliminary analyses.

Changes Between the 2011 and 2017 Grade 8 Assessments

The most noticeable change between the 2011 and 2017 administrations of the NAEP writing assessment was in the type of device students used to compose their responses (a laptop in 2011 and a tablet with attached keyboard in 2017). Differences between these devices are presented in Table 1. They include differences in the sizes of the devices, display resolution, input devices, and

Table 1. Features of the digital devices used in the NAEP writing assessment: Grade 8, 2011 and 2017

Features	2011	2017
Hardware type	13" Dell and Toshiba laptops	Microsoft Surface Pro 4 (SP4)
Display size	13" diagonal	11.5" diagonal
Display aspect ratio	16:10 widescreen	4:3 standard aspect ratio
Display resolution	1280x800, 116 pixels per inch	2160x1440, 216 pixels per inch
Input devices	Built-in keyboard, trackpad, USB mouse	Attached keyboard, trackpad, USB mouse and active stylus
Software/Interface	Custom Windows application	Custom web application

software. For example, the tutorial to instruct students on how to take the assessment was adapted to be compatible with the features of the 2017 writing device. Also, students participating in the different subject-area assessments were tested in the same classrooms in

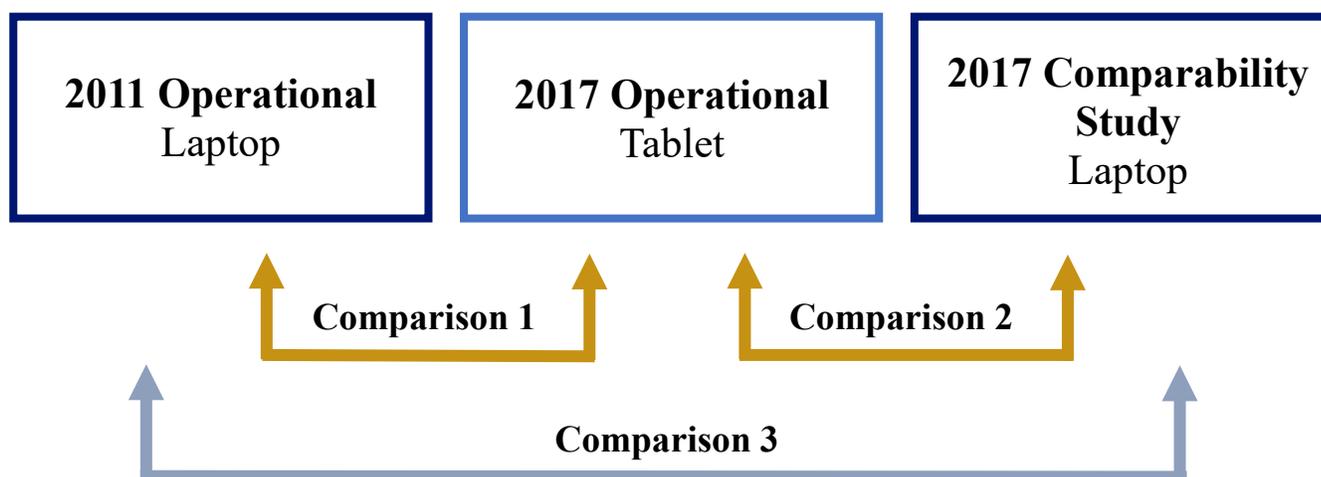
2017. In 2011, students took the writing assessment separately from students taking other NAEP assessments because writing was administered digitally, while other subjects were still being administered on paper. The format of the writing assessment itself did not change: students were asked to respond to two writing tasks and had about 30 minutes to compose and type each response.

The 2017 Writing Comparability Study

NCES conducted a comparability study in 2017 at grade 8 to determine whether changes in the digital device students used to compose their responses had a differential effect on students' writing performance. The study included about 3,000 students, used one third of the tasks common to the 2011 and 2017 operational assessments, and was administered two months after the 2017 operational assessment. This separate study provided valuable information about the effects of different devices on students' writing skills.

As figure 1 shows, the study allowed multiple comparisons of the digital writing performance of students at grade 8.

Figure 1. Performance comparisons of NAEP digital writing administration conditions: Grade 8



Due to its small scale (about 3,000 students and one third of the assessment tasks) and the time the study was conducted (two months after the administration of the assessment), this study did

not have sufficient precision and desired generalizability to be used as a *bridge* study, which NCES generally conducts to link scores from consecutive assessments in the same subject and grade when there is a major change. In 2017, NCES conducted bridge studies to link scores between the 2017 and 2015 mathematics and reading assessments because the assessments transitioned from a paper-and-pencil mode in 2015 to a digitally based mode in 2017. A bridge study was not undertaken to link scores between the 2017 and 2011 grade 8 writing assessment because the assessment had been previously administered on computer and because the existing literature did not suggest that the transition from one form of digital device to another would affect students' assessment performance.

Preliminary Findings: Grade 8

Task-level writing scores highlight aspects of the performance patterns that call for additional analysis. The three sets of comparisons conducted at grade 8 showed a pattern of lower performance for students assessed in the 2017 operational tablet condition compared to students assessed in either the 2011 laptop condition or the 2017 laptop condition.

Table 2: Preliminary findings of three comparisons: Grade 8

Comparison 1 2011 laptop vs. 2017 tablet	Comparison 2 2017 tablet vs. 2017 laptop	Comparison 3 2011 laptop vs. 2017 laptop
<ul style="list-style-type: none"> • Lower performance for 2017 tablet for all major reporting groups • Median essay length was shorter for 2017 tablet • In 2017, about 50% of students reported never using a tablet in their English language arts class in 2017. 	<ul style="list-style-type: none"> • Lower performance for tablet than laptop • Median essay length shorter for tablet than laptop • Almost all students in the tablet condition reported using a laptop in their English language arts class at least once or twice a month. 	<ul style="list-style-type: none"> • Lower performance for 2017 laptop compared to 2011 laptop

These results suggest that students' responses might have been affected by the use of a different device. NCES cannot be certain, however, that the change in performance is due solely to the change in device; other differences in administration (e.g., changes in the tutorial, assessing students separately vs. together) may also be factors.

Preliminary Findings: Grade 4

Data sources for analysis at grade 4 included the 2017 performance data, student responses to questions about their familiarity with digital devices, process data collected as part of the 2017 assessment, as well as student performance and process data collected in the 2012 pilot study.

Comparison of student performance on tablets in 2017 and on laptops in the 2012 pilot study showed similar patterns in terms of student performance, though students in 2017 wrote slightly longer responses than in 2012 and typed slightly faster. However, further analysis of the 2017 data revealed concerning results. For example, 47 percent of grade 4 students reported never using a tablet in school. Moreover, student performance showed strong relationships with how much and how fast they wrote, suggesting a strong dependence on students' keyboarding skills.

Steps Undertaken by NCES

NCES shared these findings with members of expert panels regularly consulted before the release of assessment results to assure NAEP's high standard of reliability and validity. Among these were NCES's Design and Analysis Committee and the NAEP Validity Studies Panel. NCES also consulted with the Committee on Standards, Design, and Methodology of the National Assessment Governing Board (NAGB). NCES, with the committed support of these expert panels and NAGB, decided not to release the results to avoid any possible misrepresentation of students' writing skills.

The panels also supported NCES's plan for releasing a special research report after completing further investigations. NCES, per the recommendations of these panels, plans to conduct additional studies on the roles that devices, interfaces, and typing skills might have on writing performance, and also to examine whether students' writing experience in the classroom aligns with how writing skills are assessed in NAEP.

The Forthcoming Report

The special report, a research and development report on the 2017 NAEP writing assessment, will describe the assessment design and administration conditions in detail; will examine patterns of performance by student groups, by various categories of tasks and by contextual

factors; and will include expert panel recommendations along with the results of additional investigations. The report is expected to provide insight into the future design and administration of digitally based NAEP writing assessments.