Lynne N. Kennette, DURHAM COLLEGE, lynne.kennette@durhamcollege.ca
Dawn McGuckin, DURHAM COLLEGE, dawn.mcguckin@durhamcollege.ca

Best Practice for Online Tests: How Long Do Students Actually Need?

ABSTRACT

Multiple choice tests are unlikely to disappear from formal education, partly due to the ease of large-scale administration and grading and their similarity to licensing exams in various fields (e.g., nursing). Despite post-secondary instructors' best intentions in giving students adequate time to complete multiple choice assessments, it can be difficult to judge the amount of time that is actually required by students, while attempting to maintain test integrity and minimize cheating behaviour. Further, much of the available literature on this topic focuses on students enrolled in four-year university programs, which are likely to differ from other post-secondary programs (i.e. two- and three-year diploma programs). The present study aims to quantify the amount of time students in two- and three-year programs actually used to answer multiple choice questions in a fully online, asynchronous, introduction to psychology course, as well as examine whether differences exist in the time used on two types of assessments: small quizzes with unlimited attempts and unit tests with only one attempt. Results showed that students used on average 39 seconds per question, though they used significantly more time on summative assessments (unit tests) compared to formative quizzes. These results can help guide pedagogical decisions, but it is also important to consider learner-specific characteristics which might affect how much time they use (or need) to complete multiple choice assessments.

KEYWORDS

student testing, test time, multiple choice, online testing

INTRODUCTION

Multiple choice tests are still a popular approach to assessment in higher education. Even prior to the pandemic, many instructors chose to provide students with online testing for some of their assessments, be it for low-stakes, knowledge-check quizzes, or higher-stakes tests/exams. But a valid concern has been how much time to provide students per question, given that too little time could affect student performance and increase stress and too much time might increase the risk for violations of academic integrity. The purpose of this study, which uses secondary data analysis, is to investigate students' actual test time behaviour by describing how many seconds per question students spend on average in an online introduction to psychology general education course at an Ontario college. With very little available data on this topic and none existing for this type of institution, this study bridges a gap which currently exists in the literature. This information is of particular importance so that educators can make more informed pedagogical decisions as practitioners when deciding how much time to allocate for multiple-choice assessments.

LITERATURE REVIEW

Historically, empirical interest in the amount of time students require for answering multiple choice questions began, as far as we can tell, in the early twentieth century with Ruch (1924), who had a strong interest in the parametric properties of tests, particularly in validity, and Paterson (1924) who focused on the development of alternatives to essay-type questions, including true-false, multiple-choice, and fill-in-the-blanks. As it relates to the time required to answer multiple choice questions, Ruch (1924) proposed that "Six 5-response items can be answered per minute of working time" and, "Seven 3-response items can be answered per minute of working time" (114). This corresponds to 16 minutes per 100 questions for 5-response items and 13.5 minutes per 100 questions for 3-response items (Ruch 1924, 113), or 9.6 seconds and 8.1 seconds (respectively), per item. Although reading speed and question length were not empirically reported in their investigation, the appendix shows that the questions were short, fact-retrieval question stems (e.g., "The capitol of the United States is") with equally short choices for responses (e.g., Boston, Chicago, New York, Washington, Baltimore), and Brothen (2012) calculated their average length to be 13.6 words per question.

Since this time, there have been recommendations provided by various academics, the most influential of which is McKeachie. In their book, McKeachie (2002) has recommended that students be given one minute per multiple choice question as a rule-of-thumb for university-level testing. While there is no literature cited in their book to support this recommendation and they do not relate this recommendation to a particular length of multiple choice question, we can assume that reading speed will affect test completion time as students must read the test in order to respond (Trauzettel-Klosinski, Dietz, and IReST Study Group 2012). Other guidelines have suggested that students should expect to spend approximately 30—60 seconds per multiple choice question that are on the lowest levels of Bloom's Taxonomy (i.e., understand and remember questions, such as definitions), which is typical in introductory and general education courses, such as ours (Anderson, Krathwohl, and Bloom 2001; Bloom 1956; Pear Deck Learning 2019; Sherman and Wildman 1982), but again, outside of the few studies mentioned here, there is limited empirical data available.

Aside from collegial recommendation, little has been done to empirically document how much time students actually spend answering multiple choice questions (Brothen 2012), perhaps due to the onerousness of collecting data from a test administered on paper, which, until recently, was likely the norm for classroom administration. However, with the ubiquitous use of online testing for both formative and summative assessments, these additional data are now more easily available. In spite of this, very little empirical research exists, with the exception of Brothen (2012) who examined how much time university students actually used per question in online multiple-choice tests (though there is no mention of item difficulty, only that the pool of questions were provided by the textbook publisher). Across four experiments, they examined both high-stakes exams and low-stakes quizzes as well as proctored and non-proctored assessments. Overall, they found that students used between 24.63 and 42.79 seconds per question, which falls somewhere between Ruch's (1924) estimate and McKeachie's (2002) rule-of-thumb. Of particular interest for our purposes, the students who participated in study one did not meet the standard admission criteria for the university (i.e., scores on the standardized American College Testing and high school percentile rank were too low for admission) and so were part of a special unit at the university. Brothen (2012) reported that these students used, on average, 27.45 seconds per question (compared to study two where the same question pool was used and the more traditional university students completed the questions in 24.63 seconds per question, on average). These results are particularly noteworthy because we suspect that these students will more closely mirror the characteristics of an Ontario college student than the typical university student does. This is because students bound for technical colleges like ours have

been building slightly different skills in high school by completing applied college-stream courses and thus, like Brothen's (2012) sample in study one, would not meet the requirements for university and thus might more closely resemble the high school students in this study.

Perhaps the lack of clear guidelines and recommendations in the literature stems from the numerous factors which can affect the amount of time it might take a student to answer a multiple-choice question, including reading speed and language proficiency, question difficulty (both in terms of content being tested and the linguistic characteristics of the question), the impact on final grade, test proctoring, paper or digital administration, and students' approach to the test or question (Attali 2016; Baron 2015; Brothen 2012; Ha, Marsic, and Yaneva 2017; Harik, Fineberg, and Clausser 2020; Juric 2020; Sage et al. 2019; Swanson et al. 2001; Swanson et al. 2006; Trauzettel-Klosinski, Dietz, and IReST Study Group 2012; van der Linden 2009; Wise and Kong 2005; Wise and Kuhfeld 2020). For example, students might favour speed over accuracy when responding to questions, particularly on low-stakes assessments (Attali 2016; Juric 2020; Wise and Kong 2005), given that that low-stakes assessments might be taken less seriously by some students, promoting rapid-guessing (Attali 2016; Brothen 2012). Although well-prepared students may be able to answer questions faster, they may still agonize over the correct answer on a high-stakes assessment, resulting in a slower response; alternately, under-prepared students might take longer to respond as they attempt to decide on the best guess or engage in rapid-guessing behaviour and take less time to respond (Brothen 2012; Juric 2020; Wise and Kong 2005; Wise and Kuhfeld 2020). Another consideration is that students could be using the low stakes quizzes as an opportunity to focus their learning/studying on the information which is likely to appear on the unit tests, using quizzes as learning opportunities, rather than assessments (a strategy termed "quiz to learn" by Brothen and Wambach 2001).

The present study

How much time do students spend answering a multiple-choice question? Past work examining this variable has not provided a clear answer, and none has looked at a college population such as ours, where a number of factors might differ compared to university students. In Ontario, Canada, where this research was undertaken, universities are institutions that focus mainly on academic and professional programs, where students typically undertake a four-year degree. Colleges, on the other hand, are institutions that focus on applied programs including in the trades and career programs, and students undertake one- two- or three-year credentials (certificates or diplomas). The mandate of colleges in Ontario is to provide career-oriented postsecondary education and training that helps students find and keep employment, meets the needs of employers, and supports the local community (Government of Ontario 2023). There are also different entrance requirements from high school (including minimum grade point averages and/or level of courses completed in high school) for degree and college programs. We would expect students who ultimately attend university versus college to be different, particularly in Ontario where children starting their high school education are streamed into applied, college-level courses, or theoretical university-level courses, preparing them for these distinct post-secondary journeys. As such, the two student populations can be quite different in terms of learner characteristics, academic skill set, and study habits (Bound, Lovenheim, and Turner 2010; Education Quality and Accountability Office 2018; Follwell and Andrew 2001; Kennette and Redd 2020; Kinnon 2016; People for Education 2019; Segedin 2012; Shaienks, Gluszynski, and Bayard 2008). This difference, and the lack of empirical research looking to understand this underrepresented college population, has been previously highlighted by other researchers who study the Ontario college population (e.g., Kennette and Chapman 2024; Kennette, Flynn, and Chapman 2023; Kennette and Jelenic 2023).

Given the variability in the amount of time reported in past literature, the lack of empirical reports on this topic, and the uniqueness of the college context, it is important to further investigate this issue. To this end, we wanted to examine how long, on average, college students actually take to answer one question on an online multiple-choice test and examine whether that average differs by type of assessment (low-stakes, small weekly quiz vs. higher-stakes, unit test). The typical college student may be different in important ways from the typical university student, whose data are provided in the studies cited above (for a discussion, see Weber and Kennette 2022), but our own anecdotal and personal experience administering tests (on a scantron in a classroom) suggests that the majority of students complete 50 multiple-choice questions in approximately 30 minutes (i.e., 36 seconds per question).

We undertook this empirical investigation to understand whether our sample of college students (in general education courses delivered across the college) would require, on average, a different amount of time for the online tests and quizzes, or whether their time usage aligns with the published literature focused on university students answering questions on paper. We hypothesized that most college students would generally use approximately 30–45 seconds per question across our psychology course assessments, but that they would spend more time on the low-stakes quizzes compared to unit tests. This is the result of the fact that the quizzes are formative assessments, which have no time limit, whereas, with the summative assessments, they are limited to two minutes per question, and some students may feel some pressure to complete it quickly. These questions align with Hutchings' (2000) "what is" questions of inquiry into the scholarship of teaching and learning.

METHOD

Participants

Following approval from the institution's research ethics board, data from students enrolled in fully online sections of GNED-1106 (Introduction to Psychology) taught by the researchers during a two-year span (2021–2022 and 2022–2023 academic years) were included in this retrospective data analysis. Across these sections, we obtained valid data from 519 students. General education courses span the entire 14-week semester (42 hours total) and follow the format of most courses at the college.

The researchers have been teaching general education psychology courses at this institution for over a decade and professors at this institution (as is the case in most colleges in Ontario) are expected to focus primarily on teaching (and service where appropriate), with no expectation of research. The teaching load is typically four to five courses per semester, depending on class size. The college is located in an urban area (population 190,000) close to a major metropolitan area (population 6.2 million).

General education courses, such as the one sampled for this study, are taught to students across the entire college as an elective. As such, we can expect our sample to roughly match the campus demographics from our institution, although we collected no specific sample-level demographic data due to the nature of our study (secondary use) and student privacy (faculty do not have access to the demographic profiles of their students). As an institution, our student population of 14,000 is comprised of approximately 21% international students with approximately 40% of our student population being under 21 years old and a little more than half (54%) aged between 21 and 35 years old, based on the data available from the 2021–2022 reporting year (Durham College n.d.). Students enrolled in general education courses, such as the one in the present study, would be enrolled in either a two-year diploma or a three-year advanced diploma program, as per the ministry requirements for programs requiring students to complete general education courses.

Materials and procedure

The course consisted of weekly quizzes (for weeks 2–14) as well as four unit tests. As such, the researchers each downloaded their own students' test-time data from the learning management system (LMS) and stripped all of the identifying information before combining the data into a single spreadsheet. We identified the semester and specific assessment (e.g., quiz one) and gave each individual student a consecutive participant identification number for this study. Because the question-level time is not captured by the LMS, we calculated an average amount of time per question based on the total time spent to complete the test. In terms of question length, quiz questions contained an average of 20.21 words per question while test questions contained an average of 18.08 words per question. We didn't measure difficulty explicitly, but all questions reflected the two lower levels of Bloom's Taxonomy (Understand and Remember), as would be expected in an introductory general education college course.

Data analyses

There were a total of 5,897 assessments across the 519 students for an average of 11.36 assessments per student, ranging from two to 18 data points per student. Not all students in the sample provided data for all available assessments (i.e., some students didn't complete some of the quizzes and tests). Any student with only one assessment (n = 9) was not included in the final sample. We also eliminated outliers based on any natural breaks in the data: data points where fewer than five seconds/question or more than 800 seconds/question were removed. We chose these cutoffs because they represented natural breaks in the data at both extremes of the distribution. The data described below includes the remaining data; due to the parameters identified above, sample size for each assessment varies.

Assessments varied in terms of average number of words per question and total number of questions. Test questions were also slightly more difficult than the weekly quiz questions, which is something we acknowledge in the following discussion. These characteristics, along with the average time students spent per question (and the corresponding standard deviation), are summarized in Table 1.

Table 1. Descriptive statistics for each assessment, showing the average amount of time spent per question (in seconds)

	Unit 1 Test	Unit 2 Test	Unit 3 Test	Unit 4 Test	Week 2 Quiz	Week 3 Quiz	Week 4 Quiz	Week 5 Quiz
N	477	478	489	480	300	268	317	313
Mean	54.39	59.70	57.46	49.70	44.16	33.17	32.89	37.38
SD	21.75	27.85	30.47	27.49	49.05	32.64	44.48	60.08
Minimum	9.38	7.20	8.18	5.00	7.50	10.00	5.00	10.00
Maximum	138.75	163.20	177.27	162.50	536.25	230.00	590.00	670.00

	Week 6 Quiz	Week 7 Quiz	Week 8 Quiz	Week 9 Quiz	Week 10 Quiz	Week 11 Quiz	Week 12 Quiz	Week 13 Quiz	Week 14 Quiz
N	319	443	200	309	315	280	302	311	296
Mean	26.54	24.50	30.80	31.26	28.93	22.96	33.72	29.41	24.27
SD	33.24	40.49	43.88	55.04	40.93	30.20	54.09	32.69	24.23
Minimum	5.46	6.00	20.00	6.00	7.50	8.57	8.57	5.46	7.50
Maximum	300.00	648.00	500.00	714.00	495.00	291.43	762.86	300.00	195.00

RESULTS

Overall (quizzes and tests combined), students took 38.65 seconds per question on average (SD = 40.50). Splitting the assessments to compare average time per question on the unit tests compared to the weekly quizzes, students took longer on the test questions (M = 55.32, SD = 27.34) than the quizzes (M = 30.57, SD = 43.28), despite the fact that tests, on average, contained fewer words per question than the quizzes. Because the mean number of words per question differed significantly for tests (M = 18.08, SD = 4.40) compared to quizzes (M = 20.21, SD = 7.58; $t(5895) = 11.39$, $p < .001$, $d = .31$), we conducted an ANCOVA to control for this variable. Controlling for question length, there is a significant difference between the amount of time students spent per question on average for tests and quizzes, with significantly more time being spent on test questions compared to quizzes ($F(1,5894) = 533.97$, $p < .001$, $\eta^2 = .083$). The covariate of the average number of words per question also significantly predicts the amount of time used, though the effect size is very small ($F(1,5894) = 6.82$, $p < .01$, $\eta^2 = .001$).

We also examined whether individual students demonstrated significant differences in how long they took to answer questions on tests compared to quizzes, so we conducted a paired-samples t-test. This difference was also significant ($t(485) = 22.93$, $p < .001$, $d = 1.04$, with students taking longer on tests (M = 54.72, SD = 23.12) compared to quizzes (M = 27.47, SD = 24.72). Again, this is despite test questions being shorter than quiz questions, thus taking less time to read.

DISCUSSION

We hypothesized that, overall, students would take approximately 30–45 seconds to respond to each question. Results align with this hypothesis, with students averaging 38.65seconds per question across all assessments. This is in line with previously published literature (e.g., Brothen 2012); however, our questions appear to be quite a bit shorter than those reported by Brothen (2012) whose questions ranged from an average of 35.72 to 52.59 words per question, despite reporting average times per questions which were similar to ours (ranging from 24.63 to 42.79 seconds per question). This points to the possibility that college students are different from university students, perhaps in their preparedness for academic assessments that aren't as hands-on, as is the case with multiple choice assessments, though more research is needed to explore this possibility.

It is also possible that online assessments (the present study and that of Brothen 2012), could take students either more or less time to complete than traditional paper assessments. Our results suggest that the online versions of these assessments do not require significantly more or less time than the paper versions. Multiple studies comparing print to online mediums have generally shown no

differences in reading speed or various other aspects of reading (e.g., comprehension) between reading on a screen or in print (Baron 2015; Sage et al. 2019).

We also predicted that, due to the formative nature of quizzes and the unlimited time given, students might spend more time per question on these assessments; however, that was not supported, as students actually took longer on the higher stakes test questions. This might be explained by the fact that students had unlimited attempts at the quizzes, so they went through them quickly (possibly guessing), knowing they would have another chance to improve their marks. Faster "guessing" behaviour does tend to occur more frequently on low-stakes tests, pointing to a strategy where students are motivated to complete it quickly rather than as accurately as possible (Juric 2020; Wise and Kong 2005). As such, speed can be an indicator of student motivation (Wise and Kong 2005), and it makes sense that students could be less motivated (i.e., faster) to complete low-stakes assessments, particularly if they have multiple attempts, though we do not have data to support this possible explanation. With the tests, on the other hand, this was not the case and only one attempt was allowed, meaning they took additional time to check their answers (perhaps even using prohibited sources such as the internet or their textbook). Attali (2016) also reported that students spent 63% less time on a low-stakes section of the Graduate Record Exam (GRE; a section used only for their research and consequently would not impact students' actual GRE scores), compared to the high-stakes of the actual GRE (Attali 2016). For students who took less time on this low-stakes assessment, their performance was also lower than those who took more time (Attali 2016). This might explain why students took less time on our low-stakes tests: they knew they could try again, so they simply attempted to get through it quickly. Future research could explore whether students were faster on tests because they were better prepared for higher-stakes assessments.

Limitations

In addition to the limitations inherent in the use of secondary data (e.g., unavailable demographic information), we also did not look at student performance (i.e., accuracy) on these assessments. It is possible, as suggested by Wise and Kong (2005), that students favoured speed over accuracy, at least for the quizzes, and that this trade-off resulted in faster response times on quizzes but lower accuracy and/or more frequent attempts. Future studies should examine this possibility. Another limitation of the present study is that the data collected were not actually per question due to the limitation of the learning management system; ideally, question-level data would be collected to improve accuracy and remove some of the noise in the present data. Future research should also continue monitoring changes to the amount of time students use on multiple choice assessments and investigate this question in different populations (e.g., international students, students requiring accommodations, students with English as an additional language, etc.). Many factors at the level of the individual student were not measured or controlled for in the present study, including baseline reading speed, language proficiency (including English as an additional language), ease of retrieval/student proficiency (i.e., how well they studied), number of choices/lures, the question difficulty (though all of ours were in the lower 2 levels of Blooms Taxonomy), linguistic characteristics of the question, etc. These can all affect how much time a student uses to answer a question (Ha, Marsic, and Yaneva 2017; Harik, Fineberg, and Clausser 2020; Swanson et al. 2001; Swanson et al. 2006; Trauzettel-Klosinski, Dietz, and IReST Study Group 2012; van der Linden 2009). For example, native speakers of English can read an average of 161 words per minute (+/-18) but across multiple languages, reading occurs at an average of 184 (=/-29) words per minute (Trauzettel-Klosinski, Dietz, and IReST Study Group 2012). Consequently, and assuming that test questions are no longer than around 29 words in length, this would suggest that well-prepared students would have enough time

to read and answer the multiple-choice question using McKeachie's (2002) guideline of one minute per question.

Also, although a detailed discussion of test time needs for neurodivergent students is outside the scope of this paper, interested readers can refer to Ofiesh and Hughes (2002) for a review of the literature on determining appropriate test time considerations for these students. Of course the principles of Universal Design for Learning (UDL) should continue to be considered as instructors determine how much time to allow for assessments (CAST 2018, 2024).

In addition to the risk of students accessing prohibited resources when completing an assessment online, a limitation of this study is that none of the online assessments were proctored. Brothen (2012) observed that students took more time answering questions on unproctored quizzes than on proctored ones, suspecting unauthorized use of materials (though this was not confirmed statistically). Another important consideration is the continuous development and improvement of Generative AI tools. This technology continues to develop additional abilities, including answering multiple choice questions. As such, the amount of time used by students as identified here is only appropriate to consider when students are honestly reading and answering the questions, not if outside tools for assistance; the probability of this occurring seems most likely during a proctored assessment. However, having a better understanding of how much time students generally take to respond to questions online may be used to help instructors identify those students who might be using these technologies in unauthorized ways (e.g., spending significantly less time might indicate automation or assistance from AI, while using much more time might suggest that a student is looking up answers in a textbook or more traditional search engine). Again, the times identified here are simply described to help guide instructors in their pedagogical decisions, which might require having a conversation with a student about the potential use of technology in unauthorized ways to complete these assessments. However, it is also important to remember that averages are simply starting points to help guide faculty when making estimates for the amount of time required, because most individual students are not average individually, but taken as a group, they can help to inform these pedagogical decisions.

CONCLUSIONS

How much time do college students need to answer a multiple-choice question? Our research suggests that the widely accepted estimate of 30–45 seconds is accurate, though instructors may consider slightly more time for major summative assessments such as unit tests (as opposed to low-stakes quizzes). It is also important to consider learners' unique, individual variables which might impact this estimate (academic experience, language proficiency, etc.).

This study used secondary data to describe the average amount of time that students used per multiple test question, answered online in one elective course at a community college. As such, it should not be taken as an unnuanced "answer" to the question of how much time to give students for online multiple-choice assessments. Equally important in informing these conversations and making these pedagogical decisions should be UDL and the heterogeneous nature of the student population including neurodivergence, linguistic diversity, and various other social and cognitive diversity that makes up our student population. The data presented here add to the body of data describing how much time students actually use (i.e., need) in order to complete these questions which can help instructors inform their own decisions about how much time their students are likely to require, and consequently, how much time instructors should allocate.

AUTHOR BIOGRAPHIES

Lynne N. Kennette (CAN) is a professor of psychology at Durham College and teaches various introductory courses. Her research interests focus on areas of scholarship of teaching and learning (SoTL), including Universal Design for Learning (UDL) and various other aspects of student learning and success.

Dawn McGuckin (CAN) is a general education professor at Durham College and sessional instructor at Ontario Tech University. She teaches psychology and history courses, as well as educational courses. Her research interests include the effects of social media on mental health and Universal Design for Learning (UDL).

ETHICS

This research was approved by the Research Ethics Board (REB) at Durham College.

REFERENCES

Anderson, Lorin W., David R. Krathwohl, and Benjamin S. Bloom, editor. 2001. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.

Attali, Yigal. 2016. "Effort in Low-Stakes Assessments: What Does It Take to Perform as Well as in a High-Stakes Setting?" *Educational and Psychological Measurement* 76 (6): 1045–58. https://doi.org/10.1177/0013164416634789.

Baron, Naomi S. 2015. *Words Onscreen: The Fate of Reading in a Digital World*. Oxford: Oxford University Press.

Bloom, Benjamin S., editor. 1956. *Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain.* New York: David McKay.

Bound, John, Michael F. Lovenheim, and Sarah Turner. 2010. "Why Have College Completion Rates Declined? An Analysis of Changing Student Preparation and Collegiate Resources." *American Economic Journal: Applied Economics* 2 (3): 129–57. http://doi.org/10.1257/app.2.3.129.

Brothen, Thomas. 2012. "Time Limits on Tests: Updating the 1-Minute Rule." *Teaching of Psychology* 39 (4): 288–92. https://doi.org/10.1177/0098628312456630.

Brothen, Thomas, and Cathrine Wambach. 2001. "Effective Student Use of Computerized Quizzes." *Teaching of Psychology* 28 (4): 292–4. https://doi.org/10.1207/S15328023TOP2804_10.

CAST. 2018. *Universal Design for Learning Guidelines Version 2.2*. https://udlguidelines.cast.org/more/downloads/#v2-2.

CAST. 2024. *Universal Design for Learning Guidelines Version 3.0*. https://udlguidelines.cast.org.

Durham College (n.d.). *DC Students.* Open Data at DC. Accessed June 15, 2024. https://durhamcollege.ca/about/office-of-research-services-innovation-and-entrepreneurship-orsie/institutional-research-and-planning/open-data-at-dc/our-students.

Education Quality and Accountability Office. 2018. "Ontario Student Achievement: Ontario's Provincial Secondary School Report." *Toronto, ON: Queen's Printer for Ontario*.

Follwell, Tianna, and Sam Andrew. 2001. "How to End Streaming in Ontario Schools." *Ontario 360.* https://on360.ca/wp-content/uploads/2021/05/ON360_EndingAcademicStreaming_v2.pdf.

Government of Ontario. 2023. "Framework for Programs of Instruction for Colleges of Applied Arts and Technology." *Ontario.ca*. Accessed December 20, 2024. https://www.ontario.ca/page/framework-programs-instruction-colleges-applied-arts-and-technology.

Ha, Le Anh, Goran Marsic, and Vanya Yaneva. 2017. "Predicting Item Response Time Using Linguistic Features." Paper Presented at the Timing Impact on Measurement in Education Conference, Philadelphia, PA.

Harik, Paul, Robert A. Fineberg, and Brian E. Clausser. 2020. "How Examinees Use Time: Examples from a Medical Licensing Examination." In *Integrating Timing Considerations to Improve Testing Practices*, edited by Melissa J. Margolis and Robert A. Feinberg, 173–89. New York: Routledge.

Hutchings, Pat. 2000. “Approaching the Scholarship of Teaching and Learning.” In *Opening Lines: Approaches to the Scholarship of Teaching and Learning*, edited by Pat Hutchings, 1–10. Menlo Park, CA: The Carnegie Foundation for the Advancement of Teaching.

Juric, D. P. 2020. “A History of Test Speededness: Tracing the Evolution of Theory and Practice.” In *Integrating Timing Considerations to Improve Testing Practices*, edited by Melissa J. Margolis and Robert A. Feinberg, 1–18. New York: Routledge.

Kennette, Lynne N., and Morgan Chapman. 2024. “Student and Faculty Perceptions of Ineffective Teaching Behaviours.” *The Canadian Journal for the Scholarship of Teaching and Learning* 15 (1). https://doi.org/10.5206/cjsotlrcacea.2024.1.15148.

Kennette, Lynne N., Kathleen Flynn, and Morgan Chapman. 2023. “Has the Pandemic Affected Students’ and Faculty’s Use and Perception of Universal Design for Learning?” *Currents in Teaching and Learning* 14 (2): 21–48. https://webcdn.worcester.edu/currents-in-teaching-and-learning/wp-content/uploads/sites/65/2023/01/Currents-14-02-Kennette-Flynn-Chapman-Student-and-Faculty-Use-and-Perception-of-Universal-Design-for-Learning.pdf.

Kennette, Lynne N., and Milan Jelenic. 2023. “Cheating: It Depends How You Define It.” *Canadian Perspectives on Academic Integrity 5* (2): 16–33. https://doi.org/10.11575/cpai.v5i2.75649.

Kennette, Lynne N., and Bibia R. Redd. 2020. “An Exploration of Study Habits: How Do Four-Year and Two-Year Colleges Compare?” *Innovation Abstracts XLII* (43). https://www.nisod.org/2020/11/10/xlii_43/.

Kinnon, Emily. 2016. “(In)Equity and Academic Streaming in Ontario: Effects on Students and Teachers and How to Overcome These.” *Ontario Institute for Studies in Education of the University of Toronto*. https://tspace.library.utoronto.ca/bitstream/1807/72216/1/Kinnon_Emily_R_201606_MT_MTRP.pdf.

McKeachie, Wilbert J. 2002. *Teaching Tips*. 11th ed. Boston: Houghton Mifflin.

Ofiesh, Nicole S., and Charles A. Hughes. 2002. “How Much Time?: A Review of the Literature on Extended Test Time for Postsecondary Students with Learning Disabilities.” *Journal of Postsecondary Education and Disability* 16 (1): 2–16. https://files.eric.ed.gov/fulltext/EJ875992.pdf.

Paterson, Donald G. 1924. *Preparation and Use of New-Type Examinations: A Manual for Teachers*. Yonkers-on-Hudson, NY: World Book Company.

Pear Deck Learning. 2019. “How to Determine the Best Length for Your Assessment.” *Pear Deck*. https://www.peardeck.com/blog/how-to-determine-the-best-length-for-your-assessment.

People for Education. 2019. “Roadmaps and Roadblocks: Career and Life Planning, Guidance, and Streaming in Ontario’s Schools.” Toronto: People for Education. https://peopleforeducation.ca/wp-content/uploads/2019/02/Roadmaps_roadblocks_WEB.pdf.

Ruch, George M. 1924. *The Improvement of the Written Examination*. Chicago: Scott Foresman.

Sage, Karen, Heather Augustine, Heather Shand, Kelsey Bakner, and Sarah Rayne. 2019. “Reading from Print, Computer, and Tablet: Equivalent Learning in the Digital Age.” *Education and Information Technologies* 24: 2477–502. https://doi.org/10.1007/s10639-019-09887-2.

Segedin, Lauren. 2012. “Listening to the Student Voice: Understanding the School-Related Factors That Limit Student Success.” *McGill Journal of Education* 47 (1): 93–107. http://doi.org/10.7202/1011668ar.

Shaienks, Danielle, Tomasz Gluszynski, and Justin Bayard. 2008. “Postsecondary Education, Participation and Dropping out: Differences Across University, College and Other Types of Postsecondary Institutions.” *Statistics Canada.* https://eric.ed.gov/?id=ED508178.

Sherman, Thomas M., and Terry M. Wildman. 1982. *Proven Strategies for Successful Test Taking*. Columbus, OH: Merrill.

Swanson, David B., Susan M. Case, David R. Ripkey, Brian E. Clauser, and Michael C. Holtman. 2001. “Relationships among Item Characteristics, Examinee Characteristics, and Response Times on the USMLE Step 1.” *Academic Medicine* 79: S114–S116. https://doi.org/10.1097/00001888-200110001-00038.

Swanson, David B., Karen Z. Holtzman, Karen Albee, and Brian E. Clauser. 2006. “Psychometric Characteristics and Response Times for Content-Parallel Extended-Matching and One-Best-Answer Items in Relation to Number of Options.” *Academic Medicine* 81: S52–55. https://doi.org/10.1097/01.acm.0000236518.87708.9d.

Trauzettel-Klosinski, Susanne, Klaus Dietz, and the IReST Study Group. 2012. "Standardized Assessment of Reading Performance: The New International Reading Speed Texts IReST." *Investigative Ophthalmology & Visual Science* 53 (9): 5452–61. https://doi.org/10.1167/iovs.11-8284.

van der Linden, Wim J. 2009. "Conceptual Issues in Response-Time Modeling." *Journal of Educational Measurement* 46: 247–72. https://doi.org/10.1111/j.1745-3984.2009.00080.x.

Weber, Iain, and Lynne N. Kennette. 2022. "College Student Study Habits: Initial Patterns and Implications." *Reinvention: An International Journal of Undergraduate Research* 7 (1): 6–10. https://cjur.ca/may-2022-volume-7-issue-1/.

Wise, Steven L., and Xiaojing Kong. 2005. "Response Time Effort: A New Measure of Examinee Motivation in Computer-Based Tests." *Applied Measurement in Education* 18 (2): 163–83. https://doi.org/10.1207/s15324818ame1802_2.

Wise, Steven L., and Megan R. Kuhfeld. 2020. "A Cessation of Measurement: Identifying Test Taker Disengagement Using Response Time." In *Integrating Timing Considerations to Improve Testing Practices*, edited by Melissa J. Margolis and Robert A. Feinberg, 150–64. New York: Routledge.