10-5-2017

Altering Assessment for Greater Achievement: An Evaluation of Students' Attitudes and Level of Engagement with mTuner

Lindsay Shaw
University of Windsor

Altering Assessment for Greater Achievement: An Evaluation of Students' Attitudes and Level
of Engagement with mTuner

By

Lindsay Shaw

A Thesis
Submitted to the Faculty of Graduate Studies
through Psychology
in Partial Fulfillment of the Requirements for
the Degree of Master of Arts at the
University of Windsor

Windsor, Ontario, Canada

2017

Altering Assessment for Greater Achievement: An Evaluation of Students' Attitudes and Level of Engagement with mTuner

by

Lindsay Shaw

APPROVED BY:

_____

G. Salinitri

Faculty of Education

_____

K. Lafreniere

Department of Psychology

_____

J. Singleton-Jackson, Advisor

Department of Psychology

August 28, 2017

DECLARATION OF ORIGINALITY

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

With technology at the fingertips of most undergraduate students, it has been difficult for instructors to fully engage them in the classroom, which has resulted in the creation of several innovative interventions, such as mTuner. mTuner integrates several cognitive learning strategies within an assessment, with the goal of actually enhancing learning, as opposed to just measuring it. In the current study, students' attitudes, level of engagement, and performance were evaluated, with specific explorations into academic entitlement (AE), test anxiety (TA), learning orientation (LO) and grade orientation (GO). Results indicated that students had very positive attitudes toward mTuner, especially in comparison to traditional multiple-choice paper and pencil (MCPP) formats. Additionally, students had very high performance scores on the mTuner assessment despite their limited engagement in the cognitive learning features, putting to question mTuner's facilitation of long-term learning. Implications and future recommendations of mTuner implementation in educational environments are discussed.

ACKNOWLEDGEMENTS

To my research advisor, committee members, mTuner consultant, mTuner developer, and family/friends, this project would be no where without your support and guidance.

To my research advisor, Dr. Jill Singleton-Jackson your passion and commitment to teaching and learning is admirable. You taught me the importance of taking innovative risks for the betterment of higher education! Thank you for your continued support and advice.

To my committee members, Dr. Kathy Lafreniere and Dr. Geri Salinitri, thank you for your invaluable feedback, enthusiasm, and encouragement throughout all stages of this project.

To my mTuner consultant, Elizabeth Ismaile, thank you for the countless meetings, emails and phone calls. Your knowledge on mTuner was what started this project's foundation.

To the mTuner developer, Steve Joordens, this project wouldn't exist without you, your assessment tool, and your permission! Your passion for innovation is commendable.

To my family and friends, especially Graeme Singleton, thank you for encouraging me every step of the way. Don't ever stop criticizing my academic jargon. And finally, to Mom, Dad, and Jordan, my academic and personal accomplishments are all because of you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1. INTRODUCTION

With the increased influence of technology and social media, university students face many distractions while trying to learn. With these distractions, professors are finding it more and more difficult to create an optimal learning environment. However, students are most engaged during assessment, which provides a unique opportunity to enhance, not just assess student knowledge. Student learning can be further differentiated between a surface approach to learning (SAL) and a deep approach learning (DAL). SAL is a more shallow level of learning that is often displayed through memorization techniques. Students often engage in this when they are trying to cram for finals. Consequently, they forget most of the information they initially studied after writing the examination. Conversely, DAL is a more meaningful, long-term approach that emphasizes understanding of the material as opposed to just recalling it (Marton & Säljö, 1976). This approach is often demonstrated using short answer examinations and essay formats where students can further elaborate on their knowledge. Most educational contexts strive for students to engage in DAL, yet traditional assessments, such as multiple-choice paper and pencil (MCPP) examinations fail to facilitate this approach (Tippin, Lafreniere, & Page, 2012). Unfortunately, these types of assessments are often used (Blouin et al., 2009) as the only performance measure simply due to the ease of distribution, especially with large classes. Yet, with convenience comes consequence as students may pass a course with an A, but have little actual long-term knowledge. In fact, students have been identified as "passive actors" rather than "active learners" when it comes to assessment practices (Chappuis, Stiggins, Arter & Chappuis, 2004).

Evidently, there is often a disconnect between the goals of assessment, and what is actually occurring. As such, there is a need for new assessment tools that foster greater long-term

learning for students. New educational technologies, such as mTuner, have been created in attempt to fill this gap in higher education (Pare & Joordens, 2009). mTuner is an online assessment tool that uses several different cognitive learning strategies to optimize the learning process. The current study reveals students' attitudes, level of engagement and performance on the mTuner assessment, gathering various perspectives from students regarding academic entitlement (AE), test anxiety (TA), learning orientation (LO) and grade orientation (GO). A comprehensive, yet practical list of recommendations will be available for instructors who want to implement mTuner in their own classrooms. A review of the literature will explore the aforementioned student characteristics, students' attitudes toward assessment, their engagement with assessments, and the varying cognitive learning strategies employed during assessments.

Student Characteristics

Learning- vs. Grade- Orientations (LOGO). Different students may exhibit varying academic orientations, which are distinguished by learning-oriented (LO) and grade-oriented (GO) behaviours. A student that is high in LO places greater importance on actually learning the material, whereas a student high in GO places greater importance on receiving a higher grade. Many studies do measure LO and GO as dichotomies with the expectation of students being high in one and low in the other (Beck, Rorrer-Woody, & Pierce, 1991; Frymier & Weser, 2001; Lawrence & Frymier, 2002; Williams & Frymier, 2007). However, researchers have found that it is possible to score high in both, and low in both. Therefore, instead of having two distinct groups (high LO/low GO, high GO/low LO) there are four (high LO/low GO, high GO/low LO, high GO/High LO, low GO/low LO). The latter group, low GO/low LO, is very rare in an academic setting as these individuals have little to no motivation for learning and are more likely in school for external reasons, such as parental force or socialization (Eison et al., 1986; Roedel,

Schraw & Plake, 1994). A student that is high in both LO and GO have the desire to learn with

an expectation of receiving a high grade as a result of their engagement in the learning process

(Marsden et al., 2005). Ultimately, less is known about the distinctive academic characteristics of

these two groups, which has led researchers to explore students within the two dichotomies as

opposed to the four categorizations.

For instance, Eison, Pollio and Milton (1986) found a variety of differences across the

two dichotomous groups. Individuals high in LO had more motivation, locus of control, abstract

reasoning, responsibility for actions, positive attitudes, greater studying methods, and less

frustration and test anxiety compared to their high GO counterparts. Additionally, those high in

GO often take more shortcuts to get to their desired academic outcome, which can result in more

risky behaviours, such as cheating (Rettinger & Jordan, 2005), plagiarism (Marsden, Carroll, &

Neill, 2005), and the overconsumption of stimulants, like energy drinks (Ianni & Lafreniere,

2014). Interestingly, correlations with gender have also been found, both in risky behaviour and

GO, with males being disproportionately represented in both. In contrast, females have been

linked to more learning-oriented behaviours (Marsden, Carroll, & Neill, 2005).

Furthermore, those higher in GO engage in more SAL (Gibbs & Simpson, 2002; Race,

2005) suggesting that limited long-term knowledge is being established for these students

(Tippin, Lafreniere, & Page, 2012), which is evident from their lower GPAs and test scores

(Beck, Rorrer-Woody, & Pierce, 1991; Page & Alexitch, 2003). As suggested by Tippin et al.

(2012), once these students are reaching senior level courses, where less multiple- choice formats

are used, they begin experiencing more failure, and in turn, more frustration. Unfortunately, the

limited exposure to additional assessment formats in their earlier years in university does not

prepare these high GO students for that reality. In fact, this exposure to failure can further

enhance students' GO, as their position in higher education may be in jeopardy and is dependent on receiving higher grades (Beck et al., 1991; Johnson & Beck, 1988). Such desperation for success may initiate the aforementioned academic risk-taking.

   Academic Entitlement (AE). Interestingly, students with high GO exhibit similar characteristics as the small but unique academically entitled population. An academically entitled student is an individual who demands higher grades, for reasons independent of their performance (Kopp, Zinn, Finney & Jurich, 2011). For instance, a student exhibiting high levels of AE may demand a high grade simply because they attend every class or do every reading. The theoretical foundation of AE was developed from generational narcissism, which is the idea that narcissism levels in students have been on the rise (Twenge, Konrath, Foster, Campbell, & Bushman, 2008). In a cross-temporal meta-analysis, Twenge et al. (2008) found a 30% increase of narcissism from 85 different samples of college students between 1979 and 2006. AE has also been linked to family dynamics, specifically parental expectations. Students higher in AE have reported higher academic demands from their parental figures, which in turn creates more anxiety about grades. This shifts their focus away from learning and mastering the course content and toward greater grade achievement (Greenberger, Lessard, Chen & Farruggia, 2008). Although not explicitly stated within this research, this suggests that students higher in AE may be experiencing more anxiety about their grades, thus adapting more of a grade-orientation over a long-term learning approach.

   Although the research on AE has been fairly recent, researchers have found several characteristics related to AE populations: higher levels of work avoidance, frustration/negative attitudes, and less motivation and responsibility for actions (Cain, Romanelli, & Smith, 2012). Although no research has explicitly measured AE and GO together, it is evident that the two

share similar characteristics. Individuals with high GO also exhibit less motivation, less responsibility, and more frustration and work avoidance than their high LO counterparts. Even further, narcissism, a personality type closely tied with AE students, has also been linked to more SAL strategies, like memorization (Bergman et al., 2010). These SAL strategies are also utilized more frequently by students with high GO (Gibbs & Simpson, 2002; Race, 2005).

Evidently, much of the literature highlights the negative aspects of AE. It is important to note that students should be "entitled" to dedicated faculty, well-designed curriculums, and helpful resources all aimed at providing them with the opportunity to learn (Cain, Noel, Smith & Romanelli, 2014). In order to maintain a successful education system, both students and professors need to be accountable in fulfilling and upholding their high standards (Karpen, 2014). As such, providing innovative assessment tools and resisting the status-quo of distributing MCPP assessments is one strategy an instructor can take.

Test Anxiety (TA). Students with higher GO and higher AE have reported higher levels of anxiety, though in two different constructs. Those high in GO have reported higher levels of test anxiety (Eison, Pollio & Milton, 1986) and those high in AE have reported higher levels of anxiety about their grades (Greenberger, Lessard, Chen & Farruggia, 2008). Although the latter is related to test anxiety (TA), no published research has explored TA in AE populations.

This is particularly important in terms of assessment as those with higher levels of TA have trouble encoding and storing information (Everson, Smodlaka, & Tobias, 1995; Thomas, Cassady, & Heller, 2017), and are less efficient in cue-utilization strategies (Cassady & Johnson, 2002). Researchers have suggested that these poorly implemented learning strategies are a result of poor studying habits, such as procrastination (Kalechstein, Hocevar, Zimmer, & Kalechstein,

1989) or avoidant coping strategies, such as mental disengagement (Stoeber, 2004; Thomas, Cassady, & Heller, 2017; Zeidner & Matthews, 2005). However, students experiencing this high TA may know the material just as well as their less anxious peers but have trouble utilizing the appropriate cognitive learning strategies to retrieve that information during assessment. This anxiety can be especially heightened in a testing environment that has time pressures (Plass & Hill, 1986), which is a characteristic of most assessments in higher education. Consequently, students with high TA often perform worse on assessments than their peers with lower TA (Decaro et al., 2011; Thomas, Cassady, & Heller, 2017; Zeidner & Matthews, 2005) and attribute their failure to external sources (Cassady, 2004). This cycle of external attribution continues into future assessments, and creates a negative perception of assessment, in turn creating a cycle of avoidant behaviours.

This cyclical process for high TA students can result in more academic risk taking, specifically in the use of cognitive enhancement (CE) drugs (Sattler & Wiegel, 2013). Cognitive enhancement drugs are nonprescribed medications, such as Adderall or Ritalin, that are used to enhance learning, memory, and attention (Normann and Berger, 2008). According to Racine and Forlini (2010), 3% to 11% of college students in the United States have used these drugs to enhance their performance in academics. Students with high TA may be more inclined to take these drugs over their lower TA peers because the perceived benefits in learning, memory, and attention, which in turn, could enhance their abilities to encode and retrieve information while studying and while taking an assessment. Interestingly, Ianni and Lafreniere (2014) found that in comparison to their high LO counterparts, students with high GO engage in more overconsumption of energy drinks, an additional stimulant that can aid learning and attention.

These risk-taking behaviours parallel very closely to the overconsumption of CE drugs,

providing more evidence that high GO may be linked to higher TA.

     In a more recent study, a healthier intervention for students with high TA was proposed

by Thomas, Cassady and Heller (2017). They suggested a need for education-based interventions

that not only teach high TA students about self-regulation strategies, but also how to implement

and sustain them in their own studying practices. These strategies, along with physical

management strategies (e.g. deep breathing, muscle relaxation, etc.), could create a multifaceted

intervention program for high TA students (Bradley, McCraty, Atkinson, Tomasino, Daugherty,

& Arguelles, 2010; Larson, Ramahi, Conn, Estes, & Gibellini, 2010; Thomas, Cassady & Heller,

2017). If students could control their TA during assessments, their performance on such

assessments may increase, along with their perceptions of the assessment process.

Attitudes, Engagement, and Performance

     More positive attitudes about assessment are important, as negative perceptions are

rooted in students' understanding of learning. And that understanding of learning is not universal

to all students, as evident from the previous section. As such, students' have very unique

perspectives on their preference for assessment, and how they approach taking tests (Struyven,

Dochy, & Janssens, 2005). For instance, MCPP assessments have had a long history of students'

utilizing a surface approach to learning, triggering more positive attitudes toward that assessment

type in comparison to open-ended/essay formats (Furnham, Batey & Martin, 2011; Scouller,

1998; Scouller & Prosser, 1994). This finding was particularly prevalent for students with higher

TA, suggesting that the underlying reasons for this preference was not due to the increased

learning in MCPP practices, but rather what students deemed as easier in both preparation and content (Traub & McRury, 1990; Birenbaum & Feldman, 1998).

Since SAL contradicts the overall goal of assessments, various other assessment types have been implemented in post-secondary institutions, such as peer-assessments, self-assessments, and online assessments, all of which encompass the same goal of learning, but in different forms. Online assessments have benefits and drawbacks for both instructors and students. For instructors, online assessments can be easily distributed and marked, which is especially important for large introductory courses. Yet, becoming familiar with online assessments and their software can be challenging and result in a time-consuming learning curve. This latter drawback can also be applied to students who have never taken or navigated an online assessment platform, which can heighten the stress and anxiety of tests even more (Özden, Ertürk, & Sanli, 2004). However, there is evidence to support the idea that heightened anxiety has decreased, as online activities have become everyday practice (Dermo 2009; Walker, Topping, & Rodriques, 2008). Online assessments can also be convenient for students with the flexibility of being taken at a place and time that is suitable for the student. This adaptability can also contribute to reductions in anxiety. Additionally, online assessments can provide quick and comprehensive feedback to students, allowing students to self-regulate their own knowledge of tested content (Miller, 2009). The benefits of such assessments must outweigh the drawbacks as students have reported positive attitudes toward the introduction of online assessments within their educational environments (Miller, 2009; Smith & Caruso, 2010).

Having positive attitudes toward learning is particularly important as it has been linked to greater academic achievement. More specifically, research has identified that the more students believe that an assessment will help their learning, the better they will likely do (Brown &

Hirschfeld, 2008). However, student responsibility, and in turn student engagement, was also linked to this process, suggesting that there are several working mechanisms in evaluating student performance. For instance, self-regulation has been deemed an important characteristic in successful students, as they take responsibility and self-direction in their learning (Zimmerman, 2001). In contrast, academically entitled and grade-oriented students demonstrate more negative attitudes and often take less responsibility for their academic actions, which likely contributes to their poor performance.

Flow theory. Additionally, positive attitudes can act as an intrinsic motivator for students, especially during online assessments. In fact, perceived enjoyment was outlined as one of the key measurable constructs of flow theory (Koufaris, 2002). Flow theory is more of a holistic understanding of engagement that is often defined as the "experience people feel when they act with total involvement" (Csikszentmihalyi, 1977, 1997). Despite its dated origin, flow theory has been applied to several domains such as technology (Lu, Zhou & Wang, 2009; Koufaris, 2002), organizational behaviour (Eisenberger, Jones, Stinglhamber, Shanock & Randall, 2005), education (Engeser & Rheinberg, 2008; Shernoff, Csikszentmihalyi, Schneider & Steele-Shernoff, 2003) and e-learning (Lee, 2010; Liu, Liao & Peng, 2005). The theory posits that if students are fully engaged in an assessment they will become fully absorbed in that activity, thus eliminating many of the distractions around them. This full absorption is triggered, in part, by the students' enjoyment in the activity. If students actually enjoy what they are learning about, and consequently have positive attitudes toward that learning, they will likely be more engaged. According to Lee (2010), e-learning can often bring students to this state of enjoyment as technology is often perceived as fun and pleasurable for young students.

However, like most psychological constructs, engagement is much more complex, and is also influenced by the interplay of individual knowledge and the perceived difficulty of the assessment. The flow theory suggests that students experience optimal flow when their skill level and the challenge level of the task are at their highest (Csikszentmihalyi, 2008). Students learn best and are engaged most when the assessment requires them to use their knowledge in a challenging format. Although this challenge is the most optimal learning environment, there are other learning environments that can create either anxiety or boredom within the students. First, if the assessment is too challenging and their knowledge is low, students will be too anxious to fully engage in the assessment process. And second, if the assessment is too easy and their knowledge is high, students will be too bored to fully engage in the assessment process. Therefore, assessments need to challenge students to an optimal point in which they can confidently demonstrate their knowledge of the topic at hand.

Cognitive Learning Strategies

Instructors can take advantage of students' need for engagement during assessment by utilizing it as a point of intervention that can enhance learning, not just assess it. Utilizing assessment in such a way is not typical in higher education classrooms, and thus presents a gap in our current system. Steve Joordens, an educational researcher at the University of Toronto recognized this gap and created an online assessment tool called mTuner, which integrates several cognitive learning strategies within the assessment. The multiple-choice format begins with free recall and ends with immediate feedback, embedding retrieval cues and second opportunities to answer questions for fewer marks. Figure 1 displays this pattern for each multiple-choice question that students encounter.

Figure 1.

MTuner process with cognitive learning strategies.
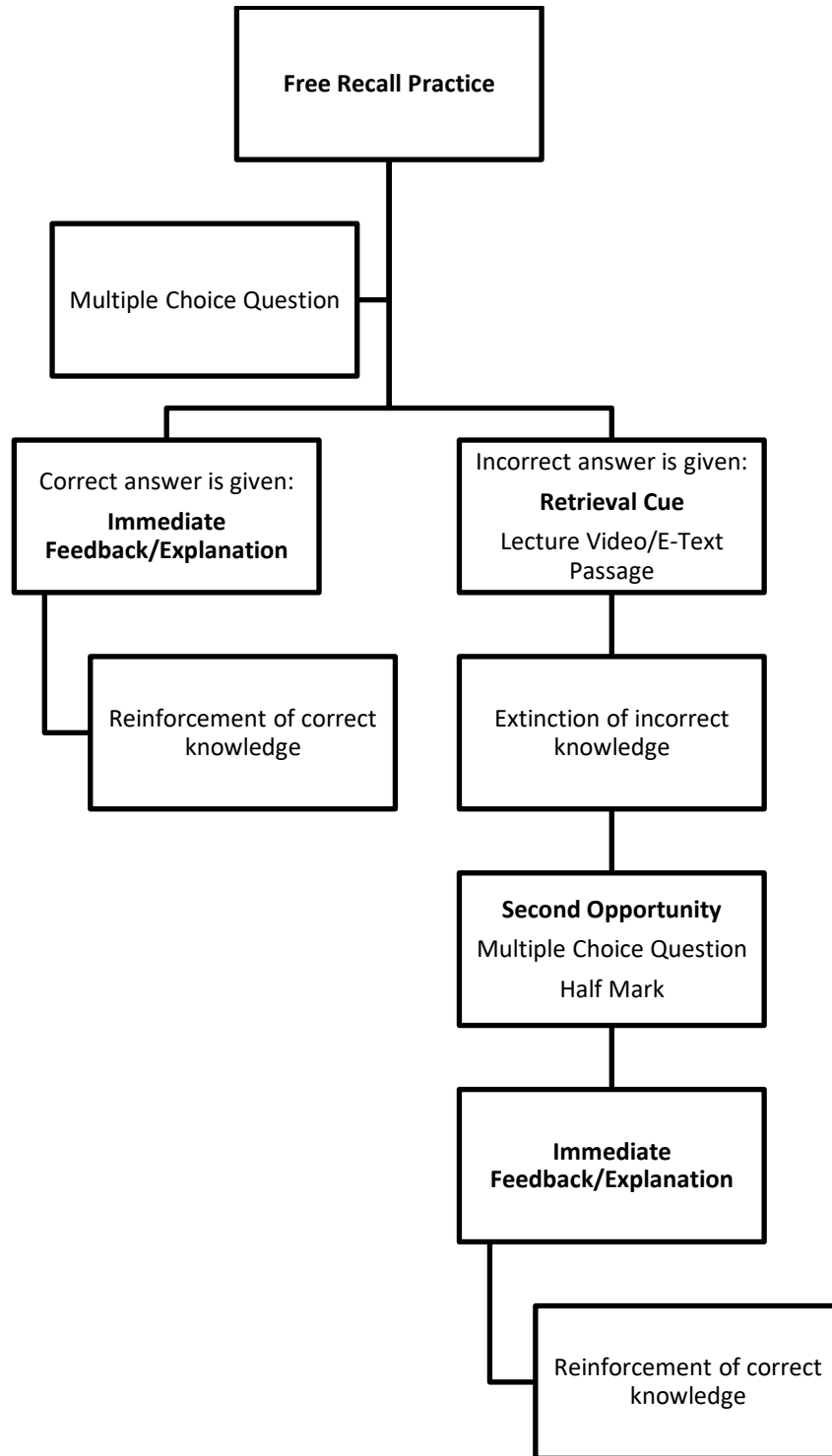
Retrieval to Free Recall. Taken together, these cognitive learning strategies enhance information retention and can strengthen long-term learning through testing, which is referred to as the testing effect (Carpenter, & DeLosh, 2006). The strength of the testing effect is dependent on the type of retrieval and the underlying mechanisms required for that retrieval. Retrieval is the process of accessing knowledge, often from our stored memory with the help of environmental retrieval cues (Karpicke, 2012). Retrieval is essential for our learning because it not only helps us access knowledge, but it also helps enhance knowledge (Roediger & Karpicke, 2006). For instance, the more we try to retrieve the same information, the more we are enhancing that knowledge, making it easier to retrieve in the future (Karpicke & Roediger, 2007, 2008; Karpicke & Zaromb, 2010). Recall tests (e.g. short answer, essay) strengthens long-term learning more so than recognition tests (e.g. multiple choice) because the retrieval process is needed for success in recall but not for success in recognition (Carpenter, & DeLosh, 2006). Students can simply "guess" the answer to a multiple-choice question through simple recognition, but this recognition is not possible during short answer assessments. Instead, students must recall information with no hints, such as those provided in multiple choice options. As students continue to practice recalling information, the retrieval routes that transfer the knowledge is strengthened (McDaniel & Masson, 1985). Students would be engaging in DAL as opposed to SAL.

During the mTuner assessment, the first feature of each multiple-choice question, requires students to engage in free recall where they can type as much information as they want in the dialogue box. This acts as a shortened essay format where students can experience DAL and prime associations. However, it is important to note that the mTuner assessment does not require students to engage in free recall, nor does it give them instructions in how they should do

so. In a study by Jordan (2012), these instructions were important for students when they were doing an online assessment that utilized an open-ended textbox. Students' answers were dependent on whether there were instructions available to them, often using more grammatical structure and writing more when a wordcount was present. Yet, in this study, students were being marked on their answers, thus encouraging them to actually engage in this process. In contrast, students are not being graded on their responses in free recall for the mTuner assessment, which in turn, may impact engagement. Despite a mark not being attached to the free recall feature, it is highly important for student learning. The more they engage in the free recall feature, the more their retrieval routes are being used and hopefully strengthened.

Unsuccessful Retrieval Attempts and Immediate Feedback. If the retrieval route is already strengthened from previous practice, successful retrieval attempts (i.e. the correct answer) are more likely to be made. This suggests that the correct information is being primed successfully in student memory. Yet, unsuccessful retrieval attempts are less understood, specifically in how they benefit learning. Often, students write MCPP examinations, receive their marks weeks later and rarely identify which items they got right and which they got wrong. Therefore, there is a greater likelihood of incorrect concepts being instilled in student memory, ultimately resulting in subsequent mistakes in later examinations (Marsh, Roediger, Bjork, & Bjork, 2007). In short, the act of choosing an answer on an exam is a form of reinforcement. Choosing an incorrect answer and never reviewing the exam to identify mistakes actually reinforces incorrect knowledge.

In a six-phase study, Kornell, Hays and Bjork (2009) found that unsuccessful retrieval actually enhanced learning when immediate feedback was present. This finding further supports the theoretical importance and multifaceted nature of the testing effect. Not only do successful

retrieval attempts enhance long-term learning, but so do unsuccessful retrieval attempts with the support of immediate feedback. There are several online assessment tools that utilize immediate feedback, such as ViLLE (Kuikka, Laakso, & Joshi, 2016; Laakso, 2010), ASSISTment (Feng, Heffernan, & Koedinger, 2009) iTest (Joglar, Martin, Colmenar, Martinez, & Hidalgo, 2010), and, of course, mTuner. In the mTuner assessment, students are given immediate feedback (i.e. explanations) for every question regardless of successful or unsuccessful attempts.

Similar to mTuner assessments, instructors have found various ways to provide feedback to their students during assessment. At times, these feedback mechanisms may not be comprehensive given the large class sizes. However, the feedback provided on the mTuner assessment differs significantly from the type of feedback that is possible on MCPP assessments, like the Immediate Feedback Assessment Technique (IFAT). The IFAT utilizes an answer-until-correct format, where students have the opportunity to scratch-off answers until they are given the correct one. In mTuner assessments, the feedback is more extensive than simply indicating what the correct answer is. Instructors can write in the feedback function as much or as little information as they would like, making reference to textbooks or other class materials. As such, students are given a more comprehensive understanding of why the given answer is right or wrong. With that, students will not further reinforce incorrect answers in their long-term memory and it is more likely that they will not make the same mistake on subsequent assessments.

Yet, the differences in feedback in the two assessment formats should not discount the IFAT. In fact, students learn more on MCPP assessments that utilize IFAT than those without any feedback at all (Dihoff, Brosvic, & Epstein, 2003; Epstein, Epstein, & Brosvic, 2001). Similarly, students had more positive attitudes toward IFAT than a typical grading scheme used in MCPP assessments (DiBattista Gissem Sinnige-Egger, Canadale & Sargeson, 2009; DiBattista

& Gosse, 2006; DiBattista et a., 2004). Having an even more comprehensive method of feedback, like the one utilized in mTuner, would likely mirror the same results. Students can self-regulate their own learning while recognizing content that they know well and content that they still need to master (Wosley, 2008). Such high-levels of self-regulation can increase engagement, motivation, and even performance (Crisp & Ward, 2008).

Although receiving immediate feedback has been beneficial for most students, past studies found that immediate feedback heightens students' test anxiety, especially when incorrect answers are made (Auerbach, 1973; Eysenck, 1982). This may be a function of the avoidant-coping strategy utilized by most students with high TA. Knowledge of their performance, especially if they are doing poorly, would negatively contribute to the on-going stressor of receiving a bad grade on the assessment. Therefore, the students with high TA would rather avoid the stress thus protecting themselves from that psychological anxiety.

Despite this making theoretical sense, more recent studies have found that using IFATs did not increase test anxiety in students (Dihoff et al., 2003, 2004), and in some cases, it actually reduced test anxiety (DiBattista & Goose, 2006). The latter may be a function of the game-like features of IFAT (DiBattista et al., 2004), or the calming nature of getting an answer correct (DiBattista & Goose, 2006). The online interactive nature of mTuner appeals to the game-like features making assessments feel far less evaluative and in turn, far less anxiety-provoking.

Second Chances & Retrieval cues. Retrieval is often cue-dependent (Tulving, 1974), meaning that students might fail to recall the information not because they did not study the material, but because they used ineffective cues to remember. Cues can come from the external or internal environment. For instance, an individual writing an introductory psychology exam

might remember they forgot to feed their dog that morning. That thought might elicit a memory

of Pavlov's classical conditioning experiment (Pavlov, 1927), aiding them in a response to a

question. If the cue helps the student further understand the concept, they can use that cue for

studying practices and future examinations. In the mTuner assessment, an incorrect response to a

multiple-choice question is followed by a retrieval cue (i.e. hint) in the form of a brief video

lecture or e-text passage. Students then have a second attempt to answer the same question for

half the mark.

This marking differs significantly from typical MCPP assessments where students are

given marks based solely on correct answers. This type of marking is typically referred to as the

number-correct grading scheme (NCGS). Yet, the NCGS only rewards students for specific

knowledge on a specific question when in reality students may have partial knowledge on a

topic. For example, they may quickly eliminate two answers as irrelevant, but struggle with two

options that have partial relevance (Dibattista, 2005). This alone demonstrates that students have

some knowledge on the topic in question, but they have yet to master it completely. In mTuner

assessments, students would be rewarded for this partial knowledge by having a second chance

to answer the question for fewer marks. Using a different online assessment format, Merrel and

colleagues (2015) found that a high percentage of students were not guessing on their second

attempt, but were actually getting it correct. They concluded that students actually learned from

their mistakes and were mastering the material. There are added benefits for including multiple

attempts within examinations, which include eliminating student anxiety and enforcing a level of

importance in student development, not just academic outcomes (Baleni, 2015).

The latter is especially important as the current higher education system promotes the

attainment of high grades and uses that as the fundamental marker of success. Learning,

especially in higher education, can not be unidimensional, with a focus solely on understanding

content. In fact, students especially in senior courses should be able to apply their content both

critically and creatively with the understanding that mistakes help transform their learning

experience. In fact, having a transcript of A's is not predictive of professional success. The

ability to learn from mistakes and adapt accordingly is a more marketable skill for a future

employee (Noll & Wilkins, 2002; Saade, Morin & Thomas, 2012). Instructors should be

modeling this philosophy in their classrooms and aligning assessments to promote that

philosophy by creating a learning-centered classroom.

mTuner Research

The theoretical foundation of mTuner has been well established, as the aforementioned

cognitive learning strategies have proven to increase learning in various educational contexts.

Yet, no published research exists on this assessment tool despite its use in post-secondary

institutions. However, Holbrook, Dupont, Power, and Joordens (2015) did conduct a pilot project

that compared mTuner to another online assessment called LEARN. The purpose of this project

was to gain an initial understanding of the students' attitudes toward the two assessments.

Students preferred mTuner over LEARN, identifying the immediate feedback feature as the most

helpful followed by the multiple attempts feature. Despite this preference, students indicated that

they did not experience greater learning in one assessment tool over the other. Although this pilot

study provides some insight into students' attitudes and performance on the mTuner assessment

there are several gaps that still need further exploring. For instance, it is important to understand

how students' attitudes compare to MCPP assessments as this is the more commonly used

assessment in large introductory classrooms. Also, understanding how students engage in the

cognitive learning features is an important indicator of their performance on the assessment.

Present Research

The current study will attempt to fill in these gaps exploring students' attitudes and level of engagement with the cognitive learning features, specifically in students with varying learning orientations, AE, and test anxiety. It is important to note that within the research questions and hypotheses the academic orientations are grouped together as if they are one unit. However, this is just for ease in communication as each component will be measured separately. Though given the similarities, it is expected that those high in anxiety will also be high in AE and high in GO and vice versa as suggested by the literature. The research questions are as follows:

1. Do students' attitudes toward mTuner differ from their attitudes toward traditional paper and pencil multiple choice examinations?

2. Will students' level of GO, AE, and test anxiety predict their level of engagement during each cognitive learning strategy (i.e. free recall, second attempt, immediate feedback, retrieval cue)?

3. Will students' attitudes and level of engagement predict their performance on the assessment?

Hypotheses. First, students' attitudes toward mTuner will be compared to their attitudes toward traditional MCPP assessment. It is expected that students will have more positive attitudes toward mTuner than MCPP. The underlying reasons for these attitudes is exploratory; however, the students higher in GO, AE and TA will likely prefer the assessment because of the features that will increase their grade with little to no effort. More specifically, these features include the option to answer the question a second time for fewer marks, and the retrieval cues (i.e. a hint). These features align closely with these students' defining characteristics of high

work avoidance (Marsden, Carroll, & Neill, 2005), limited motivation (Cain, Romanelli, &

Smith, 2012; Eison, Pollio & Milton, 1986) and greater satisfaction with surface approach

techniques (Berman et al., 2010; Tippin, Lafreniere, & Page, 2012). In contrast, it is expected

those higher in LO will prefer mTuner because of the learning-oriented features, allowing these

students to engage in DAL. These expectations are linked to their positive attitudes, motivation,

(Eison, Pollio & Milton, 1986), and feelings of empowerment in the classroom (Houser &

Frymier, 2009).

Second, student engagement will be explored within each cognitive learning strategy.

Similar to the previous hypothesis, students with higher levels of TA, GO and AE will report

lower levels of engagement in free recall and immediate feedback and higher levels of

engagement in the second attempts feature and retrieval cues. The opposite is expected in

students with higher LO.

And finally, it is expected that the more positive and the more engaged students are with

assessments the better their performance on that assessment will be (Brown & Hirschfeld, 2008;

Fredricks, Bulumenfeld, & Paris, 2004). Although previous research was conducted in MCPP

assessments, such a relationship is expected in the current study using mTuner, ultimately

demonstrating a predictive relationship.

CHAPTER II. METHODS

Sample

   Pretest. In total, 410 participants enrolled in a first-year introduction to psychology course at the University of Windsor completed the presurvey, 102 were male, 302 were female, and six chose not to specify. Given the requirement of enrollment in the introduction to psychology course, the majority of students (62.4%) were in their first year (M_{age} =19.74, SD= 3.09) and came from the arts, humanities and social sciences departments (58.1%), as well as the other science subdisciplines (21.8%) within the university. However, there were also students who had a double major (n=13) or were enrolled in Business (n=35), Human Kinetics (n=23), Nursing (n=7), Engineering (n=1), or Education (n=1). Approximately 66.5% of students identified as Caucasian or European (n=258), while 11.1% were Middle Eastern, 5.9% Asian, 5.7% African American, 4.9% East Indian, 4.6% Mixed, 1.0% Hispanic, and .3% Aboriginal/Metis/Inuit. Furthermore, the majority (61.7%) of students had high TA, low LO (57.6%), and low GO (52.7%). Despite being the majority, the dichotomous high/low cut-offs for each of these groups were fairly equal. The exception, however, was AE, as only a small portion of the sample (15.3%) consisted of high AE students.

   Post-test. Unfortunately, the study had drastic levels of attrition, as only 155 students completed the post-test survey. Yet, this smaller sample mirrored much of the presurvey sample as the majority identified as being Caucasian/European (69.1%) females (77.1%) in their first year of study (M_{age} =19.46, SD= 2.12) in the arts, humanities and social sciences departments (59.2%). Yet, many of the groupings for the student orientations changed, with the majority having low TA (62.9%), high LO (62.6%), and high GO (54.8%). The disproportionate dichotomy of AE remained the same, with only 16.7% of students being categorized as high AE.

See Tables 1 through 3 for a detailed outline of this sample's descriptive statistics. Given the

predictive nature of the research questions, all results pertaining to those questions will use this

sample.

Table 1.

Ethnic Background (Post-test)

Ethnic Background	Frequency	Percentage of Sample
Caucasian/European	103	69.1
Middle Eastern	16	10.7
Asian	9	6
African American	9	6
East Indian	6	4
Mixed	5	3.4
Hispanic	1	.7

Note. (N=149) as 6 participants chose not to specify their ethnicity.

Table 2.

Academic Major (Post-test)

Discipline	Frequency	Percentage of Sample
Arts, Humanities, Social Sciences	90	59.2
Science	31	20.4
Business	14	9.2
Human Kinetics	7	4.6
Double Major	7	4.6
Nursing	3	2.0

Note. (N=152) as 3 participants chose not to specify their academic major.

Table 3.

Student Orientations (Post-test)

Student Orientation (N)	High/Low	Frequency	Percentage of Sample
AE *	High	25	16.7
	Low	125	83.3
GO	High	85	54.8
	Low	70	45.2
LO	High	97	62.6
	Low	58	37.4
TA**	High	53	37.1
	Low	90	62.9

Note. *(N=150) as 5 participants chose not to complete the survey.
     ** (N=143) as 12 participants chose not to complete the survey.

Participant compensation. All students were recruited from the university's research pool and were given 0.5 bonus points for their participation in the presurvey, and 0.5 bonus points for their participation in the post survey. They were treated in accordance with ethical guidelines from the APA/CPA. See Appendix A for the study's consent letters. Both introduction to psychology courses were taught by the same instructor. The same content and the same evaluations were used in both sections thus ensuring consistency across the sample.

Measures

Participants completed two sets of online surveys and were given a unique ID code, which was used to match participants over the sessions. These ID codes assured that students would remain anonymous to the instructor in the course. Thus, their participation in the study, or lack thereof, would not have an influence on the instructor's feelings or perceptions about the students. Participants were also required to provide additional demographic information, specifically their gender, age, ethnicity, program, and year of study. Performance was measured using students' final grades on the exam, their final grades in the course, and their final mTuner grades. By consenting to the study, students were consenting to the release of their grades (See consent letters in Appendix A). When all final grades were submitted, the lead researcher connected the grades to the student and removed any identifying information in the dataset leaving only the unique ID codes to match the participants. The lead researcher had access to the grades because of her role as mTuner administrator and GA for the course. For ethical reasons, this role was limited to the technical administration and handling of the software having no contact with the students or any of the marking in the course.

LOGOII. Learning and grade orientations were assessed using the LOGO II scale developed by Eison et al. (1983). This scale had two sections, the first with 16 items regarding specific attitudes (e.g., I dislike extra assignments that are not graded) on a 5-point Likert-type scale (1=strongly disagree to 5=strongly agree). The second section had 16 items regarding behaviours (e.g., I cut classes when confident that lecture material will not be on the exam) on a 5-point Likert-type scale (1=never to 5=always). In both sections, the questions were equally divided, half pertaining to LO and the other half GO. The 16 learning-oriented items and 16 grade-oriented items were combined to create totaled scores in their respective orientations. Recommended cut-offs from previous studies were then used to categorize students in their high/low dichotomies. The reliability of this 32-item scale (α=.72) was consistent with prior reliability (α=.70) analyses (Levine, 2003; Purcell, 2010). See Appendix B for the full LOGOII scale.

AEQ. AE was assessed using the 8 item Academic Entitlement Questionnaire (AEQ), which is a smaller, theoretically constructed measure from the original that had 26 items (Kopp et al., 2011). It is measured using a 5-point Likert-type scale (1=strongly disagree to 5=strongly agree). As such, students' scores can range from 8 to 40, with a cut-off point of 24. Scoring above that cut-off point would result in high AE, and scoring below that cut-off point would result in low AE. As previously reported, the majority of the sample fell within the latter category (*M*=18.73, *SD*=5.39). The reliability of this 8-item scale was higher (α=.83) than that found in previous study's (Kopp et al., 2011; α=.80). See Appendix C for the full AEQ scale.

CTA. The cognitive test anxiety scale (CTA), which has 27 items, was used to measure students' test anxiety (Cassady et al., 2002). An example item was, "I lose sleep over worrying about examinations,", which was measured on a 4-point Likert-type scale (1=not at all typical of

me to 4=very typical of me). Some items had to be recoded so that they were positively weighted (See Appendix C). Possible scores ranged from 27 to 108, with higher scores indicating higher test anxiety. A cut-off point of 69 was used to dichotomize low and high levels. The reliability of this scale (α=.83) was lower than prior reliability (α=.91-93) analyses (Cassady, 2004; Furlan, Cassady, & Perez, 2009). See Appendix D for the full CTA scale.

SCoA VI. A modified version of the 33-item student conception of assessment scale (Brown, Irving, Peterson, & Hirschfeld, 2009) was used to measure students' attitudes toward the two assessments. The original scale had four dimensions: affect/social benefit, irrelevance, improvement and external attributions. The original affect/social benefit dimension measured both students' personal enjoyment as well as their perception of classroom enjoyment related to assessment experience using eight items. The number of items was reduced to three, eliminating the social aspect as its relevance is not important for the attitudinal construct in the current study. An example item was, "traditional multiple-choice assessment is an enjoyable experience to me". The irrelevance dimension measures students' negative perceptions of assessment, which originally contained seven items was reduced to three items. An example item is, "traditional multiple-choice assessments are unfair to students." These items had to be recoded so that they were positively weighted. The improvement dimension measures students' and teachers' roles in using assessment to improve upon their academic performance originally containing 10 items, and was reduced to three items. An example item is, "My teachers use traditional multiple-choice assessments to help me improve." Finally, the external attribution dimension originally measured the contributions of assessment with regards to students' future performance and the school's academic quality using eight items. This dimension was reduced to three items focusing on current and predictive learning that are a result of assessment. An example item is,

"Traditional multiple-choice assessment results show how intelligent I am." Items measuring

school's academic quality was eliminated as its relevance was not important for the altitudinal

construct in the current study. In addition to irrelevant items, the current scale was modified due

to practical barriers involving student time and motivation while completing the scales. All items

were measured using a 5-point Likert-type scale (1=strongly disagree to 5=strongly agree). In

measuring attitudes toward mTuner, the same modified scale was used, replacing the words

traditional assessment with mTuner assessment where appropriate. The reliability of this scale

was consistent with prior reliability (α=.80) analyses (Brown et al., 2009) both for the MCPP

version (α=.79) and the mTuner version (α=.85). See Appendix E for the full ScoA VI scale.

Students were also asked two additional questions. First, indicating their assessment

preference between MCPP or mTuner, with the option of indicating equal preference or no

preference at all. And second, which of the cognitive learning strategies they found most helpful

(i.e. free recall, retrieval cue, immediate feedback, second attempt). Students were forced to

answer only one option for both questions, resulting in easier categorization, and ultimately,

easier analyses.

Engagement Scale. Given the specific nature of the questions and the need to

understand learning strategies specifically, an engagement scale was created as no current scale

was relevant. However, three of the total items were adapted from the National Survey of

Student Engagement (NSSE) scale, which measures the quality of engagement of students across

universities in the United States and Canada producing an overall internal reliability of .8

(Tendhar, Culver & Burge, 2013). Additionally, all items were developed based on the

theoretical understanding of the engagement construct, which has been defined as the amount of

time students devote to a desired outcome (Kuh, 2001) as well as their level of collaboration and

communication with staff and peers (Coates, 2007). Therefore, the scale contained eight total items. Four of which measured how often students engaged in each cognitive learning strategy, and the other four measuring the level of engagement with their peers and professor with an example item being, "During the current mTuner midterm, generally how often did you read the feedback given after each multiple-choice question." The items were measured on a 5-point Likert-type scale (1=never from 5=all the time). The combined items produced a low reliability score (α=.64). Therefore, the items were used individually for all subsequent analyses. See Appendix F for the full Engagement scale.

Procedure

The current study was open from October 2016 until May 2017, falling within two academic semesters. Depending on students' program and interest in the introductory to psychology course, students could have been enrolled in both semesters or just one. Each semester students had three mTuner midterms followed by a final MCPP exam. There was also an optional mTuner practice test that students could use as a studying tool for the final exam. Prior to the first mTuner midterm, students were provided with instructions that would assist them in the assessment process. See Appendix G for these instructions. The instructor also gave detailed instructions on what the assessment was, the importance of the assessment, and how it was graded, during a lecture at the beginning of the semester.

Given the varying enrollment times, and limited first semester sign-up, the pre-survey was available for the whole duration of the study. The purpose of the pre-survey was to gather a baseline attitudinal score of students' attitudes toward MCPP assessments. However, it also contained the student characteristic measures, specifically learning and grade-orientation (LOGO

II), test anxiety (CTA) and academic entitlement (AEQ). The post-study opened in December to

give students in the first semester enough time to become familiar with mTuner and was open

until the end of the study period. The post-survey measured students' attitudes of mTuner so that

any changes in attitudes from MCPP assessments to mTuner could be analyzed. A measure on

students' engagement with mTuner was also included in the post-survey. See Table 4 for the

measures, their corresponding number of items and alpha reliabilities. Although students were

encouraged to participate in both surveys, the researcher did not follow-up with students to

remind them to do so. The researcher also had the role of mTuner administrator and had ethical

restrictions on the amount of communication with students. Although, in hindsight, there could

have been strategies put in place to eliminate the amount of attrition. As such, there were various

consequences to this procedure, which are discussed later in the limitations section.

Table 4.

Pre-test Post-test Measures

Survey	Measures	# of Items	Prior α	Current α
Pre-test				
	LOGO II	26	.70	.72
	AEQ	8	.80	.83
	CTA	27	.91	.83
	ScoA VII MCPP	12	.80	.79
Post-test				
	Engagement mTuner	8	N/A	.64
	ScoA VII mTuner	12	.80	.85

N/A indicates no prior reliability as measure was created for this study.

<div align="center">CHAPTER III. RESULTS</div>

Overview

All data sources were analyzed using SPSS software producing results for, independent samples t-tests, paired sample t-tests, and multiple regression analyses (MRA). Assumptions were checked for each test and the results of such analyses are included in their respective sections.

Exploratory Analyses of Student Characteristics

Independent sample t-tests were used to determine whether relationships existed between the differing student characteristics.

GO & TA.. The first two-part analysis explored the relationship between TA and GO. First, exploring whether students with high GO ($n=80$) differ from their low GO ($n=63$) counterparts on their level of TA. And second, whether students with high TA ($n=53$) differ from their low TA ($n=90$) counterparts on their level of GO. It was initially hypothesized that a relationship would exist between these two student characteristics as similarities have been found in previous studies. Prior to the first analysis, the assumptions were assessed to ensure there were no violations in the dataset. The groups were independent of one another and had an unequal but adequate sample size (>15 per group). Normality and the presence of outliers were assessed through histograms and boxplots respectively, meeting both of the assumptions. And finally, statistical inspection of Levene's test indicated a non-significant value (p=.87), suggesting that this assumption was also met. With all of the assumptions satisfied, the results indicated a statistically significant difference in the amount of TA in students with high GO (*M*=66.59, *SD*=10.27) and low GO (*M*=61.84, *SD*=10.59), *t*(141)=-2.71, *p<0.05*, 95% CI [1.28, 8.21]. This

statistically significant difference had a medium-sized effect, *d*=0.46. As hypothesized, students with high GO had more TA than their peers with low GO.

For the second analysis, the assumptions were again analyzed. The groups were independent of one another and had an unequal but adequate sample size (>15 per group). Normality and the presence of outliers were assessed through histograms and boxplots respectively. Although normality was met there was a presence of four outliers on the low TA variable. The analysis was run with and without the outliers. Both analyses produced similar results and therefore the outliers were kept in for the final analysis (Parke, 2013). And finally, a statistical inspection of Levene's test indicated a non-significant value (p=.38), suggesting that this assumption was also met. Similar to the previous finding, there was a statistically significant difference in the amount of GO in students with high TA (*M*=47.74, *SD*=6.81) and low TA (*M*=45.13, *SD*=7.91), *t*(141)=-2.00, *p*<0.05, 95% CI [-5.18, -.03]. This statistically significant difference had a medium-sized effect, *d*=0.36. Similar to the first analysis and aligning with the initial hypothesis, students with higher TA had more GO than their peers with lower TA.

GO & AE.. The second two-part analysis explored the relationship between GO and AE. First, exploring whether students with high GO (*n*=84) differ from their low GO (*n*=64) counterparts on their level of AE. And second, whether students with high AE (*n*=25) differ from their low AE (*n*=125) counterparts on their level of GO. Prior to the first analysis, the assumptions were assessed to ensure there were no violations in the dataset. The groups were independent of one another and had an unequal but adequate sample size (>15 per group). Normality and the presence of outliers were assessed through histograms and boxplots respectively. Although normality was met there was a presence of two outliers on the low GO variable. The analysis was run with and without the outliers. Both analyses produced similar

results and therefore the outliers were kept in for the final analysis (Parke, 2013). And finally, a

statistical inspection of Levene's test indicated a non-significant value (p=.09), suggesting that

this assumption was also met. There was a statistically significant difference in the amount of AE

in students with high GO (*M*=20.36, *SD*=5.37) and low GO (*M*=16.12, *SD*=4.38), *t*(148)=-5.19,

p<0.001, 95% CI [2.62, 5.85]. This statistically significant difference had a small-sized effect,

d=0.17. This finding matched the initial hypothesis as students with higher GO had more AE

than their lower GO peers.

For the second analysis, the assumptions were again analyzed. The groups were

independent of one another and had a very unequal but adequate sample size (>15 per group).

Normality and the presence of outliers were assessed through histograms and boxplots

respectively. Although normality was met there was a presence of eight outliers, five on the low

AE variable and three on the high AE variable. The analysis was run with and without the

outliers. Both analyses produced similar results and therefore the outliers were kept in the final

analysis (Parke, 2013). And finally, a statistical inspection of Levene's test indicated a non-

significant value (p=.98), suggesting that this assumption was also met. Similar to the first

analysis there was a statistically significant difference in the amount of GO in students with high

AE (*M*=51.92, *SD*=7.53) and low AE (*M*=44.98, *SD*=7.12), *t*(148)=--4.41, *p*<0.001, 95% CI [-

10.06, -3.83]. This statistically significant difference had a large-sized effect, *d*=0.96. Similar to

the previous finding and aligning with the initial hypothesis, students with high AE had more GO

than their peers with low AE.

GO & LO.. The third two-part analysis explored the relationship between GO and LO.

First, exploring whether students with high GO (*n*=85) differ from their low GO (*n*=66)

counterparts on their level of AE. And second, whether students with high LO (*n*=94) differ

from their low LO (*n*=51) counterparts on their level of GO. Prior to the first analysis, the

assumptions were assessed to ensure there were no violations in the dataset. The groups were

independent of one another and had an unequal but adequate sample size (>15 per group). There

was a presence of three outliers, one on the high AE variable and two on the low AE variable.

The analysis was run with and without the outliers. Both analyses produced similar results and

therefore the outliers were kept in for the final analysis (Parke, 2013). There was a positive skew

in the data but after dividing the skewness and kurtosis values by the standard error, the

departure from normality was not too extreme as it fit well within the rule of thumb of ±1.96.

This paired with the statistical test's robustness to normality violations, no changes were made to

the data. And finally, a statistical inspection of the Levene's test indicated a non-significant value

(p=.63), suggesting that this assumption was also met. With that, there was not a statistically

significant difference in the amount of LO in students with high GO (*M*=47.11, *SD*=6.41) and

low GO (*M*=46.09, *SD*=6.22), *t*(149)=-.98, *p*>0.05, 95% CI [-1.04, 3.07] despite the

hypothesized relationship from previous research.

  For the second analysis, the assumptions were again analyzed. The groups were

independent of one another and had a slightly unequal but adequate sample size (>15 per group).

Normality and the presence of outliers were assessed through histograms and boxplots

respectively. Although normality was met there was a presence of six outliers, three on the high

LO variable and three on the low LO variable. The analysis was run with and without the

outliers. Both analyses produced similar results and therefore the outliers were kept in for the

final analysis (Parke, 2013). And finally, a statistical inspection of Levene's test indicated a non-

significant value (p=.61), suggesting that this assumption was also met. Similar to the first

analysis there was not a statistically significant difference in the amount of GO in students with

high LO (*M*=45.96, *SD*=7.84) and low LO (*M*=46.50, *SD*=7.20), *t*(149)=--.42, *p*>0.05, 95% CI [-3.10, 2.02] despite the hypothesized relationship from previous research.

LO & AE. The fourth two-part analysis explored the relationship between LO and AE. First, exploring whether students with high LO (*n*=97) differ from their low LO (*n*=53) counterparts on their level of AE. And second, whether students with high AE (*n*=25) differ from their low AE (*n*=125) counterparts on their level of LO. Prior to the first analysis, the assumptions were assessed to ensure there were no violations in the dataset. The groups were independent of one another and had an unequal but adequate sample size (>15 per sample). Normality and the presence of outliers were assessed through histograms and boxplots respectively, and both assumptions were met. Similarly, the assumption of homogeneity was not violated as the Levene's test indicated a non-significant value (p=.22). As a result, there was not a statistically significant difference in the amount of AE in students with high LO (*M*=18.70, *SD*=5.67) and low LO (*M*=18.11, *SD*=4.81), *t*(148)=.64, *p*>0.05, 95% CI [-1.23, 2.41].

For the second analysis, the assumptions were again analyzed. The groups were independent of one another and had an unequal but adequate sample size (>15 per group). Normality and the presence of outliers were assessed through histograms and boxplots respectively. Although normality was met, there was a presence of two outliers, one on the low AE variable and the other on the high AE variable. The analysis was run with and without the outliers. The analyses produced different results. Keeping the outliers in resulted in a statistically significant difference in the amount of LO in students with high AE (*M*=49.12, *SD*=7.64) and low AE (*M*=46.24, *SD*=5.93), *t*(148)=-2.11, *p*<.05, 95% CI [-5.58, -.18]. However, after removing the outliers no statistically significant difference was present suggesting that the outliers had an influence on the dataset. With that, the two outliers were winsorized matching the

next highest data point. And finally, a statistical inspection of Levene's test indicated a non-significant value (p=.47), suggesting that this assumption was also met. There was not a statistically significant difference in the amount of LO in students with high AE (*M*=49.12, *SD*=7.64) and low AE (*M*=46.24, *SD*=5.93), *t*(148)=-1.92, *p*>.05, 95% CI [-5.10, .07].

LO & TA. The fifth two-part analysis explored the relationship between LO and TA. First, exploring whether students with high LO (*n*=91) differ from their low LO (*n*=52) counterparts on their level of TA. And second, whether students with high TA (*n*=53) differ from their low TA (*n*=90) counterparts on their level of LO. Prior to the first analysis, the assumptions were assessed to ensure there were no violations in the dataset. The groups were independent of one another and had an unequal, but adequate sample size (>15 per group). Normality and the presence of outliers were assessed through histograms and boxplots respectively, and both assumptions were met. And finally, a statistical inspection of Levene's test indicated a non-significant value (p=.34), suggesting that this assumption was also met. As a result, there was not a statistically significant difference in the amount of TA in students with high LO (*M*=65.27, *SD*=11.09) and low LO (*M*=63.13, *SD*=9.77), *t*(141)=1.16, *p*>0.05, 95% CI [-1.51, 5.79].

For the second analysis, the assumptions were again analyzed. The groups were independent of one another and had an unequal but adequate sample size (>15 per group). Normality and the presence of outliers were assessed through histograms and boxplots respectively. Although normality was met there was a presence of two outliers both on the low TA variable. The analysis was run with and without the outliers. Both analyses produced similar results and therefore the outliers were kept in the final analysis (Parke, 2013). And finally, a statistical inspection of Levene's test indicated a non-significant value (p=.83), suggesting that

this assumption was also met. Unlike the first analysis there was a statistically significant

difference in the amount of LO in students with high TA (*M*=48.49, *SD*=5.86) and low TA

(*M*=45.47, *SD*=6.57), *t*(141)=-2.79, *p*<.05, 95% CI [-5.16, -.88]. This statistically significant

difference had a medium-sized effect, *d*=0.48. Therefore, students with high TA had higher

levels of LO than their lower TA peers.

TA & AE. The sixth two-part analysis explored the relationship between TA and AE.

First, exploring whether students with high TA (*n*=52) differ from their low TA (*n*=90)

counterparts on their level of AE. And second, whether students with high AE differ from their

low AE counterparts on their level of TA. Prior to the first analysis, the assumptions were

assessed to ensure there were no violations in the dataset. The groups were independent of one

another and had an unequal but adequate sample size (>15 per group). Normality and the

presence of outliers were assessed through histograms and boxplots respectively. Although

normality was met there was a presence of one outlier on the low TA variable. The analysis was

run with and without the outlier. Both analyses produced similar results and therefore the outliers

were kept in the final analysis (Parke, 2013). And finally, a statistical inspection of Levene's test

indicated a non-significant value (p=.19), suggesting that this assumption was also met. There

was not a statistically significant difference in the amount of AE in students with high TA

(*M*=18.90, *SD*=5.78) and low TA (*M*=17.93, *SD*=5.00), *t*(140)=-1.05, *p*>0.05, 95% CI [-2.80,

.86] despite the hypothesized relationship from previous research.

For the second analysis, the assumptions were again analyzed. The groups were

independent of one another and had an unequal but adequate sample size (>15 per group).

Normality and the presence of outliers were assessed through histograms and boxplots

respectively. Although there were no outliers, there was a negative skew in the data but after

dividing the skewness and kurtosis values by the standard error, the departure from normality

was not too extreme as it fit well within the rule of thumb of ±1.96. This paired with the

statistical test's robustness to normality violations, no changes were made to the data. And

finally, a statistical inspection of Levene's test indicated a non-significant value (p=.34),

suggesting that this assumption was also met. Similar to the first analysis there was not a

statistically significant difference in the amount of TA in students with high AE (*M*=66.78,

SD=12.14) and low AE (*M*=63.95, *SD*=10.30), *t*(140)=-1.17, *p*>.05, 95% CI [-7.61, 1.95] despite

the hypothesized relationship from previous research.

Additional descriptive statistics including minimums, maximums, means, and standard

deviations are provided in Table 4.

Table 5.

Descriptive statistics for student characteristics.

Student Characteristics	n	Minimum	Maximum	Mean (SD)
Test Anxiety	143	38.00	86.00	64.50 (10.64)
Academic Entitlement	150	8.00	34.00	18.49(5.37)
Grade Orientation	151	24.00	72.00	46.15(7.60)
Learning Orientation	151	28.00	72.00	46.66(6.33)

Research Question 1

  Attitudes toward assessment. The majority of participants (63.9%) preferred mTuner as an assessment practice (n=99), over MCPP assessments (n=22), though some liked both assessment practices equally (n=17), or had no preference at all (n=17). The multiple cognitive learning features of mTuner likely contributed to this preference. As such, students indicated that having a second chance to answer the question for fewer marks was valued the most (37.2%), followed by having hints (29.5%), and having feedback (25.6%). Very few students found none of the features helpful (1.9%), nor did they prefer the free recall feature (5.8%).

  A paired samples t-test was used to indicate if students' attitudes changed from pre-test to post-test. More specifically, whether their baseline attitudes toward MCPP assessments differed from their attitudes toward mTuner. First, assumptions were analyzed to identify any violations within the sample. Fortunately, the sample size was adequate as there were more than 15-30 cases per group (n=155). Similarly, normality was not violated, as it was both visually inspected, through histograms and normal Q-Q plots, and statistically inspected, through skewness and kurtosis values. However, a visual inspection of box plots showed two outliers within the sample. Yet once an analysis was done with and without the outliers, there was no significant differences in the results. Given that, and the smaller sample size, the outliers were kept in. As initially hypothesized, there was a statistically significant difference in students' attitudes toward the differing assessment practices, $t(140)$=-6.09, $p<0.001$, 95% CI [-5.48, -2.79]. Students' baseline attitudes toward MCPP assessment at pre-test (M=38.83, SD=6.06) were much lower than their attitudes toward mTuner at post-test (M=42.96, SD=7.11). This indicated a statistically significant change in assessment attitudes across the academic semester. This change had a medium-sized effect, d=0.51.

Additional exploratory analyses were conducted to identify whether any of the four student characteristics differed on their attitudes toward mTuner and MCPP assessments, using independent samples t-tests. Four outliers were found. Analyses were conducted with the outliers kept in and removed. The outliers did impact the results of the analysis and were therefore winsorized, bringing their values closest to the next highest extremes. Some of the data appeared to be non-normal through visual inspection of histograms, but after dividing the skewness and kurtosis values by the standard error, the departure from normality was not too extreme as it fit well within the rule of thumb of ±1.96 limits. Similarly, this assumption is fairly robust to deviations of normality. Finally, homogeneity of variance was not violated as all of the Levene's tests were not statistically significant ($p>.05$). With that, no groups statistically significantly differed on their attitudes toward mTuner, except individuals with high TA ($M=44.84$, $SD=5.99$) and low TA ($M=42.06$, $SD=6.72$), $t(137)=-2.45$, $p<0.05$. Students with high TA had more positive attitudes toward mTuner than their low TA peers. This statistically significant difference had a medium-sized effect, $d=0.43$. Similarly, no groups statistically significantly differed on their attitudes toward MCPP assessments despite the evidence from previous research.

Research Question 2

Engagement. The first engagement construct measured the amount of time students devoted to each cognitive learning strategy on a 5-point Likert-type scale. Students "only sometimes" engaged in the cognitive learning features, specifically the immediate feedback ($M=3.28$, $SD=1.15$), retrieval cues ($M=2.95$, $SD=1.35$), and the free recall ($M=2.59$, $SD=1.26$). The limited engagement in the latter cognitive learning feature, aligns closely with students' attitudes as only 5.8% preferred this feature over the others. Yet, the highest engagement for students was the second opportunity feature where they used it almost every time ($M=3.98$,

SD=1.22). Similarly, this engagement aligns closely with student attitudes, as the second opportunity feature was preferred most by the students.

Additional questions explored students' amount of collaboration and communication with staff and peers, resulting in even lower levels of engagement than the cognitive learning features. Before the assessment, students indicated that they "almost never" provided help to their peers (*M*=2.43, *SD*=1.13), sought help from their peers (*M*=2.14, *SD*=1.12), or sought help from the professor about the mTuner assessment (*M*=1.85, *SD*=1.03). Interestingly, this lack of communication and collaboration, extended into the assessment process, with students indicating that they "almost never" worked with their peers during the mTuner assessment (*M*=2.08, *SD*=1.36). However, it is important to note that these results came from students' self-reports. Therefore, the accuracy in this collaboration may not resemble how the students actually worked during the assessment process.

Predicting engagement in immediate feedback. Next, a multiple regression analysis was used to identify whether the four student characteristics predicted engagement in the immediate feedback feature. In assessing the assumptions all were satisfied, except for the presence of outliers. The sample size was adequate with at least 15-30 participants per prediction variable. Independence of observations, multicollinearity and normality were all assessed statistically. For their respective tests, all statistics were satisfied with a Durbin-Watson value of 2.23, variance inflation factors (VIF) less than 10, and skewness and kurtosis values within the appropriate ranges (Field & Miles, 2010). Linearity, heteroscedasticity, and normality were all visually assessed through, scatterplots, boxplots, normal Q-Q plots, and histograms, all meeting their required guidelines. The presence of outliers was assessed using standardized residual scores (cut-off of 2.5), Leverage scores, and Cooks distance (less than 1). Four outliers were

identified, but after running the analysis with and without them it did not make any significant changes to the model or R^2 value. This, paired with the smaller sample size, resulted in all outliers being kept in the analysis.

The model, which contained the four predictor variables, was statistically significant, $F(4,137)=3.34$, $p<0.05$, $R^2= 0.09$, $R^2_{Adjusted}= 0.06$, though only 9% of variance was accounted for. Additionally, learning orientation ($M=46.65$, $SD=6.38$), was the only statistically significant predictor, $b =0.04$ ($SE=0.02$), $\beta=.21$, 95% CI [0.01, 0.07], $t(141) = 2.53$, $p < .05$. This indicates that an increase in LO would result in heightened engagement in the immediate feedback feature. In contrast, AE, GO and TA did not predict engagement in the immediate feedback feature. This provides partial support of the initial hypothesis of learning-oriented students having higher engagement in the immediate feedback feature while students with high TA, GO, and AE would have less engagement in the feature. The full regression model with all accompanying predictors is outlined in Table 5.

Table 6.

Regression model for engagement in immediate feedback.

Predictors	b	SE	β	t	p	R	R^2	Adjusted R^2
Academic Entitlement	-.03	.02	-.14	-1.47	.14	.30	.09	.06
Grade Orientation	-.02	.01	-.12	-1.29	.20			
Learning Orientation	.04	.02	.21	2.53	<.05			
Test Anxiety	.01	.01	.10	1.12	.27			

Predicting engagement in retrieval cues. Next, a multiple regression analysis was used to identify whether the four student characteristics predicted engagement in the retrieval cue features. In assessing the assumptions all were satisfied except for the presence of outliers. The sample size was adequate with at least 15-30 participants per prediction variable. Independence of observations, multicollinearity and normality were all assessed statistically. For their respective tests, all statistics were satisfied with a Durbin-Watson value of 2.09, variance inflation factors (VIF) less than 10, tolerance values smaller than .10, and skewness and kurtosis values within the appropriate ranges (Field & Miles, 2010). Linearity, heteroscedasticity, and normality were all visually assessed through, scatterplots, boxplots, normal Q-Q plots, and histograms, all meeting their required guidelines. The presence of outliers was assessed using standardized residual scores (cut-off of 2.5), Leverage scores, and Cooks distance (less than 1). Four outliers were identified, but after running the analysis with and without them it did not make any significant changes to the model or R^2 value. Therefore, the outliers were kept in the analysis.

The model, which contained the four predictor variables, was statistically significant $F(4,137)=5.86$, $p<0.001$, $R^2= 0.15$, $R^2_{Adjusted}= 0.12$, though only 15% of variance was accounted for. There were two statistically significant predictors, LO ($M=46.65$, $SD=6.38$), $b =0.04$ ($SE=0.02$), $\beta=.20$, 95% CI [0.01, 0.08], $t(141) = 2.39$, $p < .05$, and GO ($M=46.08$, $SD=7.62$), $b =-0.06$ ($SE=0.02$), $\beta=-.33$, 95% CI [-0.09, -0.03], $t(141) =-3.62$, $p < .001$. The two predictor variables had differing directional relationships. An increase in LO led to an increase in the retrieval cue engagement, whereas an increase in GO led to a decrease in the retrieval cue engagement. However, AE did not predict engagement in the retrieval cue feature. Though it is important to note that TA ($M=64.41$, $SD=10.63$) was approaching statistical significance, $b =0.02$

(*SE*=0.01), β=.15, 95% CI [-0.01, 0.04], *t*(141) = 1.87, *p* =.06. These findings provide partial

support for the second hypothesis. It was initially hypothesized that students with high TA, GO,

and AE would have more engagement in the retrieval cue feature as it helps contribute to a

higher grade. Additionally, it was hypothesized that students' with high LO would engage less in

the retrieval cue. However, the opposite occurred for both LO and GO students. The full

regression model with all accompanying predictors is outlined in Table 6.

Table 7.

Regression model for engagement in retrieval cues.

Predictors	b	SE	β	t	p	R	R^2	Adjusted R^2
Academic Entitlement	.00	.02	.01	.02	.98	.38	.15	.12
Grade Orientation	-.06	.02	-.33	-3.62	<.01			
Learning Orientation	.04	.02	.20	2.39	<.05			
Test Anxiety	.02	.01	.15	1.87	.06			

Predicting engagement in second chance feature. Next, a multiple regression analysis was used to identify whether the four student characteristics predicted engagement in the second chance/multiple attempts feature. In assessing the assumptions all were satisfied. There was no presence of outliers or influential observations in the sample. This was assessed using standardized residual scores (cut-off of 2.5), Leverage scores, and Cooks distance (less than 1). The sample size was adequate with at least 15-30 participants per prediction variable. Independence of observations, and multicollinearity were assessed statistically. For their respective tests, all statistics were satisfied with a Durbin-Watson value of 1.74, variance inflation factors (VIF) less than 10, tolerance values smaller than .10. Linearity, and heteroscedasticity were visually assessed through scatterplots and boxplots meeting their required guidelines. There was a slight negative skew in the data but after dividing the skewness and kurtosis values by the standard error, the departure from normality was not too extreme as it fit well within the rule of thumb of ± 1.96. This paired with the statistical test's robustness to normality violations, no changes were made to the data.

The model, which contained the four predictor variables was not statistically significant, $F(4,136)=1.33$, $p>0.05$. With that, students' AE, GO, LO and TA did not predict their level of engagement in the second chance feature not supporting the initial hypothesis. It was hypothesized that students with higher AE, GO and TA would engage more in the second attempt feature because it helps contribute to a higher grade. The opposite was expected for students with high LO as they would have been more academically prepared and would not need to use this feature as often as their peers. The full regression model with all accompanying predictors is outlined in Table 7.

Table 8.

Regression model for engagement in second chances.

Predictors	b	SE	β	t	P	R	R^2	Adjusted R^2
Academic Entitlement	-.03	.02	-.12	-1.29	.20	.19	.04	.01
Grade Orientation	-.01	.02	-.03	-.32	.75			
Learning Orientation	-.01	.02	-.07	-.83	.41			
Test Anxiety	.02	.01	.13	1.52	.13			

Predicting engagement in free recall. Next, a multiple regression analysis was used to identify whether the four student characteristics predicted engagement in the free recall feature. The sample size was adequate with at least 15-30 participations per predictor variable. A total of five outliers were found, which was assessed using standardized residual scores (cut-off of 2.5), Leverage scores, and Cooks distance (less than 1). After running the analysis with and without them it did not make any significant changes to the model or R^2 value. Therefore, all outliers were kept in the analysis. Independence of observations, and multicollinearity were assessed statistically. For their respective tests, all statistics were satisfied with a Durbin-Watson value of 2.09, and variance inflation factors (VIF) less than 10. Linearity and heteroscedasticity, were all visually assessed through, scatterplots, boxplots, normal Q-Q plots, and histograms, all meeting their required guidelines. There was a positive skew in the data but after dividing the skewness and kurtosis values by the standard error, the departure from normality was not too extreme as it fit well within the rule of thumb of ±1.96. This paired with the statistical test's robustness to normality violations, no changes were made to the data.

The model, which contained the four predictor variables, was statistically significant $F(4,137)=6.80, p<0.001, R^2= 0.17, R^2_{Adjusted}= 0.14$, though it only accounted for 17% of variance. Additionally, all four predictors were statistically significant, LO (M=46.65, SD=6.38), b =0.07 (SE=0.02), β=.33, 95% CI [0.04, 0.1], $t(141) = 4.12, p < .001$, AE (M=18.29, SD=5.30), b =0.05 (SE=0.02), β=.20, 95% CI [0.01, 0.09], $t(141)$ =2.29, $p < .05$, GO (M=46.08, SD=7.62), b =-0.03 (SE=0.02), β=-.19, 95% CI [-0.06, -0.1], $t(141) = -2.11, p < .05$, and TA (M=64.41, SD=10.63), b =-0.02 (SE=0.01), β=-.18, 95% CI [-0.04, -0.01], $t(141) = -2.20, p < .05$. The latter two predictor variables had a negative relationship, indicating that an increase in GO and TA led to a decrease in engagement in the free recall feature. In contrast, LO and AE had a positive

relationship, indicating that an increase in LO and AE led to an increase in engagement in the

free recall feature. This partially supported the initial hypothesis. It was initially hypothesized

that students with high LO would have more engagement in free recall and GO, TA and AE

would engage less in this feature. The full regression model with all accompanying predictors is

outlined in Table 8.

Table 9.

Regression model for engagement in free recall.

Predictors	b	SE	β	t	p	R	R^2	Adjusted R^2
Academic Entitlement	.05	.02	.20	2.29	<.05	.41	.17	.14
Grade Orientation	-.03	.02	-.19	-2.11	<.05			
Learning Orientation	.07	.02	.33	4.12	<.01			
Test Anxiety	-.02	.01	-.18	-2.20	<.05			

Research Question 3

Performance. Although the sample was typical of most university classrooms, the final grades were higher (M=79.79, SD=9.30), which was likely a function of extremely high mTuner grades (M=90.60, SD=8.91), balanced with average MCPP final exam scores, (M=65.40, SD= 13.00).

Differences in performance. A paired samples t-test was used to identify whether the difference between the final MCPP grades and final mTuner grades was statistically significant. First, assumptions were analyzed to identify any violations within the sample. Fortunately, the sample size was adequate as there were more than 30 cases per group (n=155). The presence of outliers was visually inspected using boxplots, four outliers and two influential observations were found. All of the outliers and influential observations were present on the mTuner grades variable except for one outlier on the final MCPP variable. The data was analyzed both with and without the outliers and influential observations. Although both analyses produce similar results, the removal of the outliers and influential observations fixed some of the normality issues. Therefore, the outliers and influential observations were winsorized to meet the second highest data point. A positive skew was still present through visual and statistical inspections. Since the test is typically robust to normality violations and the integrity of the data wanted to be kept, the data was not transformed. As expected the difference between the two assessment practices was statistically significant, t(121)=21.51, $p<0.001$, 95% CI [23.06, 27.73], and had a large-sized effect, d=.88. With that, students performed significantly better on mTuner (M=90.83, SD=7.88) than on the MCPP assessment (M=65.44, SD= 12.89).

Predicting final mTuner scores. An additional multiple regression analysis was conducted to determine whether engagement and attitudes toward mTuner predicted mTuner grades. Independence of observations, and multicollinearity were assessed statistically. For their respective tests, all statistics were satisfied with a Durbin-Watson value of 1.67, and variance inflation factors (VIF) less than 10. Similarly, there were no presence of outliers or influential observations in the dataset which was assessed using standardized residual scores (cut-off of 2.5), Leverage scores, and Cooks distance (less than 1). However, there were several other violations of the assumptions in this analysis. As expected, the final mTuner grades were positively skewed, which impacted the linearity and homoscedasticity of the data. Although transformations with reflections could be used to help alleviate these issues, it may also create the presence of new problems and it will change the integrity of the data. With that, bootstrapping with 1000 samples was used. Regardless of this technique, the model was not statistically significant, $F(5,113)=1.06$, $p=.25$, $R^2= 0.06$, $R^2_{Adjusted}= 0.01$, nor were any of the predictor variables. Therefore, attitudes toward mTuner and engagement in the four cognitive learning features did not statistically predict final mTuner grades ($M=90.60$, SD=8.91). This finding did not fit within the initial hypothesis, which hypothesized that positive attitudes and heightened engagement would result in better performance. The full regression model with all accompanying predictors is outlined in Table 9.

Table 10.

Regression model for mTuner performance.

Predictors	b	SE	β	t	p	R	R^2	Adjusted R^2
Attitudes mTuner	.23	.14	.15	1.67	.10	.24	.06	.01
Engage Feedback	.75	.88	.09	.85	.40			
Engage Hint	.98	.63	-.15	-1.55	.12			
Engage Second Chance	-.54	.79	-.06	-.68	.50			
Engage Free Recall	.03	.64	.01	.05	.96			

Predicting final exam scores. Finally, a multiple regression analysis was used to determine whether final mTuner grades, engagement and attitudes toward both assessment practices predicted final exam scores. The sample size was adequate with at least 15 participants per prediction variable. There was no presence of outliers or influential observations within the dataset which was assessed using standardized residual scores (cut-off of 2.5), Leverage scores and Cooks distance (less than 1). Independence of observations, and multicollinearity were statistically inspected, and met, with a Durbin-Watson statistic of 1.90 and tolerance and VIF levels in their appropriate range. Homoscedasticity, normality and linearity were visually inspected using scatterplots and histograms. Although the latter was met, there were some issues with normality and heteroscedasticity. Normality was negative skewed and there was a slight fanning pattern on the standardized plot. Bootstrapping of 1000 samples was used as this technique can help alleviate the aforementioned problems.

The model, which contained the seven predictor variables, was statistically significant $F(7,105)=3.78$, $p<0.05$, $R^2= 0.20$, $R^2_{\text{Adjusted}}= 0.15$, though only 20% of variance was accounted for. Additionally, attitudes toward MCPP assessments (M=39.07, SD=6.45), b =0.54 (SE=0.19), β=.27, 95% CI [0.20, 0.88], $t(112) = 2.87$, $p < .01$, and final mTuner grades (M=39.07, SD=6.45), b =0.45 (SE=0.15), β=.28, 95% CI [0.20, 0.88], $t(112) = 3.07$, $p < .01$, were the only statistically significant predictors. Both predictor variables had a positive relationship. With that, as students' attitudes toward MCPP increased, their final exam scores increased as well. Similarly, as students' mTuner grades increased, their final exam scores increased as well. In contrast, attitudes toward mTuner, and engagement in the four cognitive learning strategies did not have a statistically significant impact in predicting final exam scores. This provided only partial support for the initial hypothesis, which hypothesized that positive attitudes and

heightened engagement would result in better performance. The full regression model with all

accompanying predictors is outlined in Table 10.

Table 11.

Regression model for final exam scores.

Predictors	b	*SE*	β	t	p	R	R^2	Adjusted R^2
Final mTuner	.45	.16	.28	3.07	<.05	.45	.20	.15
Attitudes MCPP	.54	.17	.27	2.87	<.05			
Attitudes mTuner	-.04	.22	-.02	-.22	.83			
Engage Feedback	1.75	1.41	.14	1.39	.17			
Engage Hint	-.78	.98	-.08	-.87	.39			
Engage Second Chance	-.86	1.20	-.07	-.77	.44			
Engage Free Recall	-.86	.84	-.08	-.94	.35			

CHAPTER IV. DISCUSSION

Technology, often literally at the fingertips of most young adults, can help students decrease their reliance on memorization and equip them with the necessary cognitive learning strategies to succeed in the classroom. As such, mTuner, like most new innovative assessments, was specifically developed to meet the growing needs of this incoming generation by steering them away from the passive nature of traditional MCPP assessments while moving them toward being more active, responsible learners. Although mTuner has a strong theoretical background that supports learning, it was important to identify whether different academically oriented students actually liked the assessment, whether they were actually engaged in the assessment, and whether they actually learned from the assessment. Explorations into student characteristics was also important to pursue in the hope that it would, in part, fill in the various gaps within the current literature.

Attitudes toward Assessment

In the current study, students had very positive attitudes toward mTuner, preferring this new assessment tool over MCPP assessments. This was evident from the attitudinal change from students' lower baseline attitudes toward MCPP assessments at pre-test (M=38.83, SD=6.06) to their higher attitudes toward mTuner at post-test (M=42.96, SD=7.11). This supports the proposed hypothesis which was based on previous research findings suggesting that technology provides an exciting and enjoyable learning outlet for students (Lee, 2010; Miller, 2009; Smith & Caruso, 2010).

It was particularly important for students with higher TA, as they had more positive attitudes than their lower TA counterparts. Having a textbook readily available and second

chances to answer questions during the assessment could be a very important safety net for

individuals who suffer from high TA. This is especially important since these high TA students

are typically less efficient in cue-utilization strategies (Cassady & Johnson, 2002) and have

trouble encoding and storing information (Everson et al., 1995). mTuner allows high TA students

to work on these cognitive learning strategies in an environment that is less anxiety-provoking

(e.g. comfort of their own home) than a traditional MCPP setting. Even further, mTuner may

appeal to these students given the game-like features that are highly interactive, which may

eliminate some of the evaluative stress of MCPP formats (DiBattista et al., 2004).

In general, students showed equal preference for all of the cognitive learning features in

mTuner, with the exception of the free recall feature. This finding is not surprising as this feature

is the most taxing on the students. It requires them to type information related to each multiple-

choice question in an open dialogue box. In doing so, students do not have access to the four

response options. Therefore, this task is the most difficult as students do not have any guidance

on what the answer might be, relying solely on their own memory. Since students knew that the

assessment was a multiple-choice format, they likely did not prepare for short answer questions

or free recall, but instead relied on recognition from memorization techniques. Despite the

underlying reasons for students' attitudes, the purpose of this free recall act, though seemingly

strenuous for no grade attached to it, has a purpose, like all cognitive learning strategies. As

such, it required engagement on behalf of the students.

Engagement

Engagement was low for the students, in their communication and collaboration with the

professor and peers. Students were instructed not to work with one another during the

assessment, which is likely why peer collaboration during assessment was low. However, this

study did rely heavily on self-reports and thus students may have lied about working together.

Additionally, the required navigation through mTuner may not have been a huge learning curve

for students, reducing the need to seek advice from their professor and peers. The more striking

findings were students' limited engagement in the four cognitive learning features.

Always prompted: free recall and immediate feedback. The way mTuner is structured,

each student would have the opportunity to engage in free recall and immediate feedback for

every single multiple-choice question. Thus, if students were utilizing the cognitive learning

strategies optimally, then they would have identified that they used these two features "all of the

time." Instead, the majority "only sometimes" used these features. This suggests that students

were, at times, quickly skipping the free recall, which was their first prompt, and quickly

skipping the immediate feedback, which was their last prompt. Skipping the free recall feature

was likely a function of three things. First, the free recall feature was the most strenuous task for

the students during the assessment and they likely did not want to put in the extra effort to

engage in it. Second, they may not have understood the importance of their engagement on the

learning process. And third, they may not have understood what to do with the box in the first

place especially since there was no explicit directions on the mTuner assessment.

Avoiding engagement: GO and TA. In fact, as students' level of GO and TA increased,

their level of engagement in free recall decreased. Students with higher GO are known to engage

in a more SAL (Tippin, Lafreniere, & Page, 2012). Since free recall takes a DAL, it makes sense

that students with high GO are more inclined to skip this feature. Additionally, time constraints

are a very significant contributor to students' TA (Plass & Hill, 1986). In the mTuner assessment

a time clock is counting down in the top righthand corner of the screen, but instead of marking

the total time of three hours, it marks the time for each section, which is two minutes. Although this clock is supposed to act as an aid for students, it could be a distraction for students with high TA, making them more inclined to skip the free recall feature and go right to the multiple-choice options.

Enhanced engagement: LO and AE. While engagement in free recall decreased for those two constructs, it increased with greater levels of LO and AE. Since high LO students have more motivation, are more focused on learning content, and take a DAL, it makes sense that they are more engaged in free recall, in comparison to their lower LO counterparts. Yet, an increase in engagement for higher AE students is less theoretically sound. It was hypothesized that students with high AE would have less engagement in free recall because their characteristics align closely with high GO students: more work avoidance, frustration/negative attitudes, and less motivation and responsibility for actions (Cain, Romanelli, & Smith, 2012). However, they also engage in more counterproductive research behaviours (Taylor, Bailey & Barber, 2015), specifically careless responding. This, paired with the limited number of high AE students, could contribute to this odd finding.

Students with higher LO also had higher levels of engagement in the immediate feedback feature. Therefore, these students actually took advantage of the explanations given to them and likely utilized the information to get a deeper understanding of the material. This engagement in DAL is typical of high LO students (Gibbs & Simpson, 2002; Race, 2005).

Not always prompted: second opportunity and retrieval cues. Although free recall and immediate feedback are present for each multiple-choice question, the retrieval cue and second opportunity feature are not always prompted. They are only prompted when students get

their first attempt at the question wrong. With that, students may never actually engage in these two cognitive learning features. Thus, having lower levels of self-reported engagement in these features would not be as problematic as the lack of engagement in free recall and immediate feedback. However, the second opportunity feature had the highest level of engagement as students used it "almost every time." Once students answered a question incorrectly the first time, they tried again for fewer marks. None of the student characteristics predicted student engagement in the second opportunity feature suggesting that they all engaged in it equally. It is understandable that students took advantage of this cognitive learning feature regardless of their orientation because they could still receive partial marks toward their final grade. Although this feature has grade-related benefits it also has learning benefits that some students may have valued. For instance, the repetitiveness and on-going retrieval needed in this feature helps strengthen student knowledge and future retrieval processes. Therefore, the second opportunity feature can appeal to all students regardless of their academic orientation.

Since the second opportunity feature is paired with a retrieval cue, students also should have engaged in this feature "almost every time". Instead, students indicated that they read or watched the hint that was prompted after the incorrect answer "only sometimes." This suggests that at times students are getting the answer wrong, skipping the retrieval cue and answering the question again. In these cases, students may have been struggling between two multiple choice options that had partial relevance. Once they realized that the first option was incorrect, they knew that their second option was the right answer, eliminating the need for a hint. Using a different assessment tool, Merrel et al. (2015) found similar results with students continuously getting their second attempt correct. They concluded that students actually learned from their mistakes and were mastering the material. However, the retrieval cues provided in mTuner are

quite comprehensive and can help the students further understand the concept. Relearning the concept through the retrieval cues and the immediate feedback can help strengthen students' future retrieval. Therefore, it was important for the students regardless if they knew the right answer the second time to actually take the effort to learn more about the concept.

Avoiding engagement: GO. Such avoidance of the retrieval cue was found in students with high GO. As students' level of GO increased, their level of engagement in the retrieval cue decreased. It was initially hypothesized that students' high in GO would engage more in the retrieval cue because it acts as a hint, and therefore it can contribute to a higher grade. However, the retrieval cue requires students to take the time to read a portion of their textbook or watch a small video lecture. Such tasks are easy to avoid for GO students, that have more work avoidance and little motivation (Eison et al., 1986).

Enhanced engagement: LO. In contrast, as students' level of LO increased as did their engagement in the retrieval cues. Students with high LO are actually taking advantage of the retrieval cue and using it to get a deeper understanding of the material.

Potential confounds to engagement. Although student characteristics can act as an indicator for the limited engagement, there are likely other factors. For instance, students may have already known the answer to the multiple-choice and, in turn, found it counterproductive to engage in some of the cognitive learning strategies like free recall and immediate feedback. Additionally, students may be unaware of the educational and cognitive benefits of going through each of the features, which is not only a flaw in mTuner implementation, but for all assessment practices. The focus, quite often, is on getting that final grade, rather than understanding what the process takes to get to that final grade. In other words, students perceive

the grade as that final marker of success as opposed to the final marker being learning and the

learning process. This again, demonstrates that disconnect between what the goals of

assessments are, active engagement, and what is actually occurring, passive recognition.

Students skipping the cognitive learning strategies in order to simply answer the question for a

correct or incorrect response is not molding them into the active learners that is the goal of this

type of assessment.

Performance

Students did extremely well on their mTuner exams with an average of 90.6%, which

differed significantly from the MCPP final exam average of 65.40%. Therefore, their high marks

on the mTuner assessment likely inhibited them to ask for help, provide help, or seek advice

from their peers or professor. It created a false sense of security, but no deep processing was

occurring because they were skipping many of the cognitive learning features.

Student attitudes and level of engagement with mTuner did not predict their mTuner

performance, despite those strong predictive relationships in previous studies investigating

MCPP assessments. The extremely high average on mTuner exams contributed to this finding, as

most students regardless of their background, attitude, or level of engagement did well on the

assessments. Such finding may be a function of flow. Students may have experienced a state of

relaxation during mTuner as the assessment task was not challenging enough. Although this

relaxation may have been beneficial for students with high test anxiety, it did not require full

engagement of the students, and thus did not support optimal learning (Csikszentmihalyi, 2008).

In contrast, this predictive relationship did exist for MCPP final exam scores. In parallel

with previous research, students' attitudes toward MCPP assessments did predict their final exam

scores. This suggests that if students perceive an assessment positively, it, to some extent, will enhance their performance on that assessment. Similarly, students' performance on mTuner midterms predicted their performance on the final MCPP exams, despite the tests being taken at different times, in different settings, with different weights toward their final grade. This relationship could be a function of learning, or it could have been influenced from other confounding variables, like pre-existing knowledge in the subject matter. Since mTuner midterm grades predicted final exam scores, it would be expected the mTuner engagement would do the same. If they were more engaged, then they would likely be learning more and taking a deep approach to learning. However, this relationship was not present. In other words, students' engagement in mTuner did not predict final MCPP exam scores. This points to a potential flaw in the mTuner assessment tool, putting to question whether students actually learned anything. For instance, the high mTuner grades could have been a result of students using their textbook, or the internet, or even working together while doing the online assessment. Although this is seemingly problematic, from an instructor's it might not be, as the goal is to learn even if that means looking it up during the assessment. Working through these problems in this way should be strengthening the learning process through DAL, as opposed to simply memorizing and then recognizing, which is typical of MCPP tests. Additionally, if students were actually engaging in the cognitive learning strategies that too should contribute to DAL. However, the problem is, if students were truly engaging in DAL, then that learning should have been translated into higher grades in the MCPP final exam, but it did not.

Yet, as outlined in the previous section, students were not optimally engaging in the cognitive learning features. This would have been a major contributor to their lower final exam grades. For instance, Smith (2007) found that the students who frequently read their feedback on

midterm assessments actually performed better on the summative assessment than their peers who never took the time to read it. Unfortunately, many of the students in the current study fell within that latter group, "only sometimes" reading the immediate feedback. As such, their final grades on the MCPP assessment suffered.

Student Characteristics

In addition to performance, attitudes, and engagement, explorations into student characteristics was also important in the hope that it would, in part, fill in the various gaps within the current literature. As expected from previous research (Eison, Pollio & Milton, 1986), a relationship between TA and GO did exist, with students' high in GO having more TA, and students higher in TA having more GO. This relationship was particularly evident in their parallel responding to free recall and immediate feedback, contributing an important issue in these students' behaviours during assessment. Yet, knowing that students high in GO also have more TA and vice versa can provide further evidence to other facets of academic behaviour, like academic risk-taking and the overconsumption of stimulants for enhanced learning, memory, and retention (Ianni & Lafreniere, 2014; Sattler & Wiegel, 2013). Students with high TA also had greater LO thus exhibiting both LO and GO characteristics, suggesting that these students may have a strong desire to learn, but are often overcome with the stress of receiving a bad grade. Knowing that the desire to learn is there can be an essential point of intervention where students can learn strategies to help them cope during assessment. Having these coping strategies readily available to them may reduce their academic risk-taking behaviours.

Despite high TA students' relationship to LO and GO, they did not differ on their level of AE, when compared to their peers with low TA. Similarly, those high in AE did not differ on

their level of TA. Although previous research found that AE students have higher anxiety, these anxiety-levels were provoked by grades (Greenberger, Lessard, Chen & Farruggia, 2008), not the assessment itself, suggesting that academic anxiety can be seen in many distinct forms.

However, a positive relationship was found between AE and GO with students' high in AE having more GO and those high in GO having more AE. The two constructs share the same academic characteristics of less motivation and responsibility, and more frustration and work avoidance (Cain, Romanelli, & Smith, 2012; Eison, Pollio & Milton, 1986). As expected, there was no relationship with LO and AE. This aligns closely with the AE construct, which is based solely on achieving high grades not actually engaging in the learning process.

Most surprisingly, no relationship was found between LO and GO students despite the expected dichotomies (Beck, Rorrer-Woody, & Pierce, 1991; Frymier & Weser, 2001; Lawrence & Frymier, 2002; Williams & Frymier, 2007). The categorization into the four groups (High LO/High GO, High LO/Low GO, Low LO/High GO, Low LO/Low GO) may provide greater insight into the distinct academic characteristics of these groups. This may have been beneficial in the current study as high TA students and high AE students fit within the High LO/High GO group. It is also important to note that varying LO students did not differ on their level of AE or TA.

Limitations & Future Research

There were several limitations to the current study, many of which were flaws in the methodological approach. First and foremost, attrition was extremely high from the pre-test to the post-test. This, in part, was a function of the distinct population of students who were currently enrolled in an introduction to psychology course that was using the mTuner

assessment. Since only one instructor was currently implementing the assessment tool, the number of participants that could be recruited was limited. The pre-test/post-test design was initially chosen for two reasons. First, students needed to have enough exposure to mTuner to answer the engagement questions. And second, the attitudinal scales for MCPP and mTuner were identical, and therefore exposure to one scale may have prompted bias in the response to the next. In hindsight, the scales could have been counterbalanced, eliminating much of the ordering effects. A larger sample size would have increased the statistical power in some of the regression analyses assuring greater confidence in the conclusions that were made and potentially uncover more statistically significant findings (Field, 2009)

Additionally, engagement could have been measured in a more comprehensive manner, relying less on self-reports and more on the detailed diagnostic output from mTuner. Not only does this detailed report indicate when the students start and finish each question, but it also provides reaction times for each cognitive learning feature. Also, all content that the students write in the free recall textbox can be viewed. In future research, free recall wordcounts, and reaction times can be used as additional forms of engagement. Similarly, a program evaluation model, with researchers actually viewing students take the assessment would also provide a more comprehensive look at this engagement process during mTuner. Having this comprehensive view would provide further, and stronger evidence to the relationships that were found. Further, many of the regression models accounted for a limited amount of variance, suggesting that other factors are important in the predictive relationship for the outcomes of student engagement and performance. For instance, constructs such as student motivation, self-efficacy, and interest in course content should be included in future studies given their expected relationship.

Finally, focus groups were initially proposed for the current investigation, but were not conducted due to the researcher's time restraints. Although not an essential piece, focus groups would have provided a more detailed account of the students' experiences. These details may have shed more light on why students had such positive attitudes toward mTuner, yet very limited engagement.

CHAPTER V. MTUNER IMPLICATIONS AND RECOMMENDATIONS

Despite the aforementioned limitations, the current research still has very important implications, especially for the future implementation of the assessment tool. The most striking finding was the high mTuner grades, despite the limited engagement in the cognitive learning strategies. It puts into question whether or not students actually learned using the tool, especially since they performed significantly worse on their MCPP final exam. Though if students, in fact, did not learn, it may not be a function of the assessment itself, but rather their lack of engagement in the tool. Students, very quickly, become accustomed to the type of learning needed for success in classroom environments. With that, students very quickly recognize that taking a SAL, with little effort and little interaction with professors, can be adequate for success both in the classroom and during assessments. Implementing a new innovative assessment tool, that integrates more of a DAL, and requires actual behavioural change from students, can be a very difficult adjustment for students. Just as this current study demonstrated, it is much easier for students to resist the change, resist the engagement, and resist the learning and instead find an easier, quicker way to receive a high grade. In the case of mTuner, that meant skipping the cognitive learning strategies, like free recall and immediate feedback, just to get to the multiple-choice options.

Recommendation 1: Educate students on the cognitive learning strategies.

Cognitive learning strategies are there for a reason. Free recall, repetitive attempts, retrieval cues, and immediate feedback all have a strong theoretical background in enhancing long-term learning. If mTuner is implemented in future academic classrooms, it is recommended that instructors or support staff alike, educate the students on the cognitive learning strategies. In

fact, Oosterhof and colleagues (2008) found that academic dishonesty in students actually

decreased when instructors took the time to communicate with them the purpose of assessment,

especially when grade-related information (i.e. rubrics) was provided. This could be particularly

important for students with higher GO and AE that engage in more academic risk taking.

Communicating this educational piece about mTuner and its corresponding cognitive learning

features, would be an easy transition into the content that students already learn in their lectures

on cognition. Additionally, the real-life application of this content can further motivate student

learning (Gikandi, Morrow, & Davis, 2011; Kember, Ho, & Hong, 2008), and even enhance

performance.

This education piece also creates a more student-centered learning environment that

promotes more student collaboration, and instructor-student interaction. These aforementioned

factors are important for building self-determination within students. According to the theory of

self-determination (Ryan & Deci, 2000), if the three psychological needs of relatedness,

competency, and autonomy are met within a specific task, then individuals will not only

experience greater psychological well-being, but they will be more motivated to achieve that

desired task. With that, it is important that instructors communicate with their students about the

mTuner assessment process, outlining their expectations of students, while showing that they

care about their learning and development. The latter can facilitate a sense of relatedness in the

classroom, while expectations can give students a sense of control and responsibility in their own

learning. And finally, integrating the cognitive learning strategies in course content can enhance

student competencies in that subject matter. Ultimately, if done effectively, simple

communication and integration of the mTuner assessment can fulfill the three psychological

needs of self-determination, ideally enhancing student motivation along the way.

Recommendation 2: mTuner as a studying tool, not just formal assessment.

The high mTuner grades can, in part, be contributed by the second opportunity cognitive learning feature, which, when prompted, rewarded students half a mark for a correct answer. These midterm averages increased the final course marks, despite the grades not being congruent to the final MCPP grades. This is a problematic indicator of student learning and can also pose problems for instructors meeting the institutional requirement of class averages for their course. As an alternative, instructors can provide mTuner as a studying tool for students, as opposed to a graded assessment. Despite the long-term learning benefits from practicing the cognitive learning strategies, students rarely have a formalized outlet for doing so. Instead, students are responsible for developing their own studying methods, which can be a grueling task for first year students who are not accustomed to assessments in higher education. Even further, it can help students with high GO and TA who are known to have greater difficulty in fostering strong and effective studying methods (Cassady & Johnson, 2002; Eison et al., 1986; Everson et al., 1995; Thomas, Cassady & Heller, 2017).

Recommendation 3: mTuner not just in higher education.

mTuner may have fallen victim to the sample of undergraduate students, many of which have poor habits embedded within their learning processes. Students in this study seem to have found ways to achieve higher grades while not actually taking a deep, meaningful approach to learning. With that, mTuner may benefit younger students. In fact, a study by Lipowski and Pyc (2014) found that young elementary school children benefitted from test-plus-restudy tasks, where they were continuously tested and were able to restudy in between with retrieval cues. Not only did they benefit in terms of performance, but they also recognized that this way of learning

was better for their memory. Additionally, mTuner's interactive interface and easy functionality would make an easier transition to younger students. Being prepared at a younger age for higher-education assessments and the importance of taking a DAL could reduce the amount of students that have poor habits entering university in their first year.

Recommendation 4. Utilize diagnostic output to mold student development

Unlike other forms of assessment, mTuner provides a very detailed diagnostic output for each individual student. Instructors have access to the students' grades, how many items they got correct on the first attempt, how many items they got correct on the second attempt, and how many items they got wrong. Even further, instructors can not only see when students started and finished the assessment, but also when they started and finished each individual question. Yet, the most descriptive piece of information is students' reaction times to each stage of the assessment. These reaction times can provide insight to specific items that students had difficult understanding, which can help mold student development in different content areas. See Figure 2 for an example of the diagnostic output.

Figure 2.

Diagnostic output from mTuner.

CHAPTER VI. CONCLUSION

To conclude, despite mTuner's limitations, the assessment tool should not be completely removed from an instructor's pedagogical toolbox. It was particularly beneficial for students with high LO and high TA, providing an outlet that harnessed their strengths and improved upon their weaknesses. However, the cognitive learning features are only meaningful when students actually take the time to use them to optimize their learning. As such, mTuner may have fallen victim to a cohort of students who are too focused on achieving that final grade and willing to bypass learning features that will help them get there. More research is needed to identify whether this type of behaviour exists in other forms of online assessment. And if so, interventions on the importance of learning and cognition should be available to students.

REFERENCES

Auerbach, S.M. (1973). Effects of orienting instructions, feedback information and trait-anxiety

level on state-anxiety. *Psychological Reports, 33,* 779-786.

Baleni, Z. (2015). Online formative assessment in higher education: its pros and cons. *The*

Electronic Journal of e-Learning, 13(40, 228-236.

Beck, H.P., Rorrer-Woody, S., & Pierce, L. (1991). The relations of learning and grade

orientations to academic performance. *Teaching of Psychology, 18*(1), 35-37.

Bergman, J.Z., Westerman, J.W., & Daly, J.P. (2010). Narcissim in management education.

Academic of Management Learning and Education, 9(1), 119-131.

Birenbaum, M., & Feldman, R.A. (1998). Relationships between learning patterns and attitudes

towards two assessment formats. *Educational Research, 40*(1), 90-97.

Blouin, R.A., Riffee, W.H., Robinson, E.T., Beck, D.E., Green, C., Joyner, P.U., Persky, A.M.,

& Pollack, G.M. (2009). Roles of innovation in education delivery. *American Journal of*

Pharmacy Education, 73(8), Article 154.

Bradley, R., McCraty, R., Atkinson, M., Tomasino, D., Daugherty, D., & Arguelles, L. (2010).

Emotion self-regulation, psychophysiological coherence, and test anxiety: Results from

an experiment using electrophysiological measures. *Applied Psychophysiology and*

Biofeedback, 35, 261–283.

Brown, G.T.L. & Hirschfeld, G.H.F. (2008). Students' conceptions of assessment: Links to

outcomes. *Assessment in Education: Principles, Policy & Practices, 15*(1), 3-17.

Brown, G. T. L., Irving, S. E., Peterson, E. R., & Hirschfeld, G. H. F. (2009). Use of interactive

Informal assessment practices: New Zealand secondary students' conceptions of

assessment. *Learning & Instruction, 19*(2), 97-111.

Brown, G. T. (2011). Self-regulation of assessment beliefs and attitudes: A review of the

students'conceptions of assessment inventory. *Educational Psychology: An International

Journal of Experimental Educational Psychology, 31*(6), 731-748.

Cain, J., Romanelli, F., & Smith, K.M. (2012). Academic entitlement in pharmacy education.

American Journal of Pharmaceutical Education, 76(10), 1-8.

Carpenter, S.K., & DeLosh, E.L. (2006). Impoverished cue support enhances subsequent

retention: Support for the elaborative retrieval explanation of the testing effect. *Memory

& Cognition, 34,* 268-276.

Cassady, J.C., & Johnson, R.E. (2002). Cognitive test anxiety and academic performance.

Contemporary Educational   Psychology, 27(2), 270 295.

Chappuis, S., Stiggins, R., Arter, J., & Chappuis, J. 2004. *Assessment for learning: An action

guide for school leaders*. Portland, OR: Assessment Training Institute.

Csikszentmihalyi, M. (1977). Beyond boredom and anxiety. San Francisco: Jossey-Bass.

Csikszentmihalyi, M. (1997). Finding flow: The psychology of engagement with everyday life.

New York: Basic Books

Coates, H. (2007). A model of online and general campus-based student engagement. *Assessment

and Evaluation in Higher Education, 32*(2), 121-141.

Crisp, V., & Ward, C. (2008). The development of a formative scenario-based computer assisted

assessment tool in psychology for teachers: the PePCAA project. *Computers &*

Education, 50(4), 1509-1526.

DeCaro, M. S., Thomas, R. D., Albert, N. B., & Beilock, S. L. (2011). Choking under pressure: Multiple routes to skill failure. Journal of Experimental Psychology: General, 140(3), 390.

DiBattista, D. (2005). The immediate feedback assessment technique: A learner-centered multiple-choice response form. *The Canadian Journal of Higher Education, 35*(4), 111 131.

DiBattista, D., Gosse, L., Sinnige-Egger, J., Candale, B., & Sargeson, K. (2009). Grading scheme, test difficulty, and the immediate feedback assessment technique. *The Journal of Experimental Education, 77*(4), 311-336.

DiBattista, D., & Gosse, L. (2006). Test anxiety and the immediate feedback assessmet technique. *Journal of Experimental Education, 74,* 311-327.

DiBattista, D., Mitterer, J.O., & Gosse, L. (2004). Acceptance by undergraduates of the immediate feedback assessment technique for multiple-choice testing. *Teaching in Higher Education, 9,* 17-28.

Dihoff, R.E., Brosvic, G.M., & Epstein, M.L. (2003). The role of feedback during academic testing: The delay retention effect revisted. *The Psychological Record, 53,* 533-548.

Dihoff, R.E., Brosvic, G.M., Epstein, M.L., & Cook, M.J. (2004). Provision of feedback during preparation for academic testing: Learning is enhanced by immediate but nto delayed feedback. *The Psychological Record, 54,* 207-231.

Eisenberger, R., Jones, J.R., Stinglhamber, F., Shanock, L., & Randall, A.T. (2005). Flow experiences at work: For high needs achievers alone? *Journal of Organizational*

Behaviour, 26, 755-775.

Eison, J.A., Pollio, H.R., & Milton, O. (1986). Educational and personal characteristics of four

different types of learning- and grade-oriented students. *Contemporary Educational*

Psychology, 11, 54-67.

Eison, J.A., Pollio, H.R., & Milton, O. (1983). Manual for use with LOGO-II. *Theory and*

Practice in Teacher Education and Other Works. Retrieved from:

http://trace.tennessee.edu/utk_theopubs/2.

Engeser, S., & Rheinberg, F. (2008). Flow, performance and moderators of challenge-skill

balance. *Motivation & Emotion, 32,* 158-172.

Epstein, M.L., & Brosvic, G.M. (2002). Students prefer the immediate feedback assessment

technique. *Psychological Reports, 88,* 889-894.

Epstein, M.L., Lazarus, A.D., Calvano, T.B., Matthews, K.A., Hendel, R.A., Epstein, B.B., et al.

(2002). Immediate feedback assessment technique promotes learning and corrects

inaccurate first responses. *The Psychological Record, 52,* 187-201.

Everson, H.T., Smodlake, I., & Tobias, S. (1995). Exploring the relationship of test anxiety and

metacognition on reading test performance: A cognitive analysis. *Anxiety, Stress and*

Coping, 7, 85-96.

Exeter, D.J., Ameratunga, S., Ratima, M., Morton, S., Dickson, M., Hsu, D., & Jackson, R.

(2010). Student engagement in very large& classes: the teachers' perspective. *Studies in*

Higher Education, 35(7), 761-775.

Eysenck, M.W. (1982). *Attention and arousal: Cognition and performance.* Berlin: Springer

Verlag.

Feng, M., Heffernan, N.T., & Koedinger, K.R. (2009). Addressing the assessment challenge with

an online system that tutors as it asses. *User Modeling and User-Adapted Interaction:*

The Journal of Personalization Research, 19(3), 243-266.

Fishbein, M., & Ajzen, I. (Eds.). (1975). Belief, attitude, intention, and behavior: An introduction

to theory and research. MA: Addison-Wesley.

Fredricks, J.A., Blumenfeld, P.C., Paris, A.H. (2004). School engagement: Potential of the

concept, state of the evidence. *Review of Educational Research, 74*(1), 59-109.

Frymier, A.B., Shulman, G.M., & Houser, M.L. (1996). The development of a learner

empowerment measure. *Communication Education, 45,* 181-199.

Frymier, A.B., & Weser, B. (2001). The role of student predispositions on student expectations

for instructor communication behaviour. *Communication Education, 50,* 314-326.

Furnham, A., Batey, M., & Martin, N. (2011). How would you like to be evaluated? The

correlates of students' preferences for assessment methods. *Personality and Individual*

Differences, 50, 259-263.

Gikandi, J.W., Morrow, D., & Davis, N.E. (2011). Online formative assessment in higher

education: A review of the literature. *Computers & Education, 57,* 2333-2351.

Greenberger, E., Lessard, J., Chen, C., & Farruggia, S.P. (2008). Self-entitled college students:

Contributions of personality, parenting and motivational factors. *Journal of Youth and*

*Adolescence, 37,*1193-1204.

Holbrook, J., Dupont, C., Power, M., & Joordens, S. (2015). *Enhancing learning through*

testing. Presented at the OND Centre for Teaching Excellence Conference, Waterloo, Ontario. Retrieved from https://uwaterloo.ca/centre-for-teaching excellence /sites/ca.centre-for-teaching   excellence/files/uploads/files/mtuner_ond_2015_final.pdf.

Houser, M.L. (2006). Do student expectations matter? Expectancy violations of instructor communication as predictors of motivation and learning: A comparison of traditional and non-traditional students. *Communication Quarterly, 54,* 331-339.

Houser, M.L., & Frymier, A.B. (2009). The role of student characteristics and teacher behaviours in students' learner empowerment. *Communication Education, 58*(1), 35-53.

Ianni, P.A., & Lafreniere, K.D. (2014). Personality and motivational correlates of energy drink consumption an misuse among female undergraduate students. *Personality and Individual Differences, 69,* 110-114.

Joglar, N., Martin, D., Colmenar, J., Martinez, I., & Hidalgo, J. (2010). ITest: Online assessment and self-assessment in mathematics. *Interactive Technology and Smart Education, 7*(3), 154-167.

Johnson, B.G., & Beck, H.P. (1988). Strict and lenient grading scale: How do they affect the performance of college students with high and low SAT scores? *Teaching of Psychology, 15,* 127-131.

Jordan, S. (2012). Student engagement with assessment and feedback: Some lessons from short answer free-text e-assessment questions. *Computers & Education, 58,* 818-834.

Kalechstein, P., Hocevar, D., Zimmer, J.W., & Kalechstein, M. (1989). Procrastination over test preparation and test anxiety. In R. Schwarzer, H. M. van der Ploeg, & G. D. Spielberger (Eds.), Advances in test anxiety research. Vol. 6. (pp. 63–76). Lisse, The Netherlands:

Swets & Zeitlinger.

Karpicke, J.D. (2012). Retrieval-based learning: Active retrieval promotes meaningful learning. *Current Directions in Psychological Science, 21*(3), 157-163.

Karpicke, J.D., & Roediger, H.L. (2007). Repeated retrieval during learning is the key to long term retention. *Journal of Memory and Lanuguage, 57,* 151-162.

Karpicke, J.D., & Roediger, H.L. (2008). The critical importance of retrieval for learning. *Science, 319,* 966-968.

Karpicke, J.D., & Zaromb, F.M. (2010). Retrieval mode distinguishes the testing effect from the generation effect. *Journal of Memory and Language, 62,* 227-239.

Kember, D., Ho, A., & Hong, C. (2008). The importance of establishing relevance in motivating student learning. *Active Learning in Higher Education, 9,* 249-263.

Kopp, J.P., Zinn, T.E., Finney, S.J., & Jurich, D.P. (2011). The development and evaluation of the academic entitlement questionnaire. *Measuring and Evaluation in Counseling and Development, 44*(2), 105-129.

Kornell, N., Hays, M.J., & Bjork, R.A. (2009). Unsuccessful retrieval attempts enhance subsequent learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(4), 989-998.

Koufaris, M. (2002). Applying the technology acceptance model and flow theory to online consumer behaviour. *Information Systems Research, 13*(2), 205-223.

Kuh, G.D., & Hu, S. (2001). The effects of student faculty interaction in the 1990s. *Review of Higher Education, 24*(30), 309-332.

Kuikka, M., Laakso, M.J., & Joshi, M. (2016). The effect of the immediate feedback by the

collaborative education tool ViLLE on learning for business mathematics in higher

education. *Journal of Education Technology Systems, 45*(1), 34-49.

Larson, H., Ramahi, M., Conn, S., Estes, L., & Gibellini, A. (2010). Reducing test anxiety

among third grade students through the implementation of relaxation techniques. *Journal

of School Counseling, 8*, 1–19.

Lawrence, K.R., & Frymier, A.B. (2002). *The relationship of student-centered teaching with

learning and grade orientation.* Paper presented at the annual meeting of the National

Communication Association, New Orleans, LA.

Lee, M.C. (2010). Explaining and predicting users' continuance intention toward e-learning: An

extension of the expectation-confirmation model. *Computers & Education, 54,* 506-516.

Levine, N. B. (2003). A multidimensional scaling analysis of perceptions of academic dishonesty

among learning- and grade-oriented students. *Dissertation Abstracts International*,

64(4), 1175.

Lipowski, S.L., & Pyc, M.A. (2014). Establishing and explaining the testing effect in free recall

for young children. *Developmental Psychology, 50*(4), 994-1000.

Liu, S.H., Liao, H.L., & Peng, C.J. (2005). Applying the technology acceptance model and flow

theory to online e-learning users' acceptance behaviour. *Issues in Information Systems,

5*(2), 175-181.

Lu, Y., Zhou, T., & Wang, B. (2009). Exploring Chinese users' acceptance of instant messaging

using the theory of planned behaviour, the technology acceptance model, and the flow

theory. *Computers in Human Behaviour, 25,* 29-39.

Marriott, P., & Lau, A. (2008). The use of on-line summative assessment in an undergraduate financial accounting course. *Journal of Accounting Education, 26*(2), 73-90.

Marsden, H., Carroll, M., & Neill, J.T. (2005). Who cheats at university? A self-report study of dishonest academic behaviours in a sample of Australian university students. *Australian Journal of Psychology, 57*(1), 1-10.

Marsh, E.J., Roediger, H.L, Bjork, R.A., & Bjork, E.L. (2007). Memorial consequences of multiple-choice testing. *Psychonomic Bulletin & Review, 14,* 194-199.

Marton, F. and Säljö, R. (1976) On qualitative differences in learning. 1 – outcome and process. *British Journal of Educational Psychology*, 46, pp. 4–11.

McDaniel, M.A., & Masson, M.E.J. (1985). Altering memory representations through retrieval. *Journal of Experimental Psychology: Learning, Memory and Cognition, 11,* 371-385.

Merrel, J.D., Cirillo, P.F., Schwartz, P.M., Webb, J.A. (2015). Multiple-choice testing using immediate feedback-assessment technique forms: Second-chance guessing vs. second chance learning. *Higher Education Studies, 5*(5), 50-55.

Miller, T. (2009). Formative computer-based assessment in higher education: The effectiveness of feedback in supporting student learning. *Assessment & Evaluation in Higher Education, 34*(2), 181-192.

Mussawy, S.A.J. (2009). *Assessment practices: Student's and teacher's perceptions of classroom assessment* (Unpublished doctoral dissertation). Univrsity of Massachusetts Amherst, MA.

Noll, C.L., & Wilkins, M. (2002) Critical skills of IS professionals: A model for curriculum

development. *Journal of Information Technology Eucation, 1*(3), 144-154.

Norman, C. & Berger, M. (2008). Neuroenhancement: Status quo and perspectives. *European*

Archives o Psychiatry and Clinical Nerusocience, 258, 110-114.

Oosterhof, A., Conrad, R.M., & Ely, D.P. (2008). *Assessing learners online.* New Jersey:

Pearson.

Özden, M.Y., Ertürk, I., & Sanli, R. (2004). Students' perceptions of online assessment: a case

study. *Journal of Distance Education, 19*(2), 77-92.

Page, S., & Alexitch, L. R. (2003). Learning- and grade-orientation, sex, and prediction of self

reported academic performance. *Psychological Reports, 92*, 320-324.

Pare, D.E., Joordens, S. (2009). The invisible researcher: Using educational technologies as

research tools for education. *Systemics, Cybernetics and Informatics, 7*(5), 37-42.

Pavlov, I.P. (1927). Conditioned Reflexes. Oxford University Press, London.

Plass, J.A., & Hill, K.T. (1986). Children's achievement strategies and test performance: The

role of time pressure, evaluation anxiety, and sex. *Developmental Psychology, 22,* 31-36.

Purcell, J.M. (2010). Learning- and grade-orientations of community college students:

Implications for instruction. *Community College Journal of Research and Practice, 34*,

497-511.

Racine, E. & Forlini, C. (2010). Cognitive enhancement, lifestyle choice of misuse of

prescription drugs? *Neuroethics, 3*, 1–4.

Rettinger, D.A., & Jordan, A.E. (2005). The relations among religion, motivation, and college

cheating: A natural experiment. *Ethics & Behaviour, 15*(2), 107-129.

Roediger, H.L., & Karpicke, J.D. (2006a). The power of testing memory: Basic research and

implications for educational practice. *Perspectives on Psychological Science, 1,* 181-210.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic

motivation, social development, and well-being. *American Psychologist, 55*(1), 68-78.

doi: 10.1037110003-066X.55.1.68

Ryan, R. M., & Deci, E. L. (2013). Toward a social psychology of assimilation: Self

determination theory in cognitive development and education. In B. W. Sokol, F. M. E.

Grouzet, U. Muller (Eds.), *Self-regulation and autonomy: Social and developmental

dimensions of human conduct* (pp. 191-207). Cambridge, England: Cambridge University

Press.

Saade, R.G., Morin, D., & Thomas, J.D. (2012). Critical thinking in e-learning environments.

Computers in Human Behaviour, 28, 1608-1617.

Sambell, K., McDowell, L. & Brown, S. (1997) 'But is it fair?': An exploratory study of

students perceptions of the consequential validity of assessment. *Studies in Educational

Evaluation, 23*(4) 349-371.

Sattler, S., & Wiegel, C. (2013). Cognitive test anxiety and cognitive enhancement: The

influence of students' worries on their use of performance-enhancing drugs. *Substance

Use & Misuse, 48,* 2230-232.

Scouller, K. (1998). The influence of assessment method on students' learning approaches:

Multiple choice question examination versus assignment essay. *Higher Education, 35*,

453–472.

Scouller, K., & Prosser, M. (1994). Student's experiences in studying for multiple choice

question examinations. *Studies in Higher Education, 19*, 267–279.

Smith, G. (2007). How does student performance on formative assessment relate to

learning assessed by exams? *Journal of College Science Teaching, 36*(7), 28-34.

Smith, S.D., & Caruso, J.B. (2010).The ECAR study of undergraduate students and information

technology. ECAR Research Study 6, Center for Applied Research, Boulder: CO.

Stöeber, J. (2004). Dimensions of test anxiety: Relations to ways of coping with pre-exam

anxiety and uncertainty. *Anxiety, Stress & Coping, 17*(3), 213–226.

Struyven, K., Dochy, F., & Janssens, S. (2005). Students' perceptions about evaluation and

assessment in higher education: A review. *Assessment and Evaluation in Higher

Education, 30*, 325–341.

Taylor, J.M., Bailey, S.F., & Barber, L.K. (2015). Academic entitlement and counterproductive

research behaviour. *Personality and Individual Differences, 85,* 13-18.

Tendhar, C., Culver, S.M., & Burge, P.L. (2013). Validating the national survey of student

engagement (NSSE) at a research-intensive university. *Journal of Education and

Training Studies, 1*(1), 182-193.

Thomas, C.L., Cassady, J.C., & Heller, M.L. (2017). The influence of emotional intelligence,

cognitive test anxiety, and coping strategies on undergraduate academic performance.

Learning and Individual Differences, 55, 40-48.

Tippin, G.K., Lafreniere, K., & Page, S. (2012). Student perception of academic grading:

personality, academic orientation and effort. *Active Learning in Higher Education, 13*(1),

51-61.

Traub, R.E. & MacRury, K. (1990). Multiple-choice vs. free response in the testing of

scholastic achievement, in: K. Ingenkamp & R.S. Jager (Eds) *Test und tens* 8: *jahrbuch*

der pädagogischen diagnostik (Weinheim und Base, Beltz Verlag), 128–159.

Tulving, E. (1974). Cue-dependent forgetting. *American Scientist, 62*(1), 74-82.

Twenge, J.M. (2001). Changes in women's assertiveness in response to status and roles: A cross

temporal meta-analysis, 1931-1993. *Journal of Personality and Social Psychology, 81*,

133-145.

Twenge, J.M., Konrath, S., Foster, J.D., Campbell, W.K., & Bushman, B.J. (2008). Egos

inflating over time: A cross-temporal meta-analysis of the narcissistic personality

inventory. *Journal of Personality and Social Psychology, 76*(4), 875-902.

Wen, M.L., & Tsai, C. (2006). University students' perceptions of and attitudes toward online

peer assessment. *Higher Education, 51*, 27-44.

Williams, K.D., & Frymier, A.B. (2007). The relationship between student educational

orientation and motives for out-of-class communication. *Communication Research*

Reports, 24, 249-256.

Wosley, T. (2008). Efficacy of instructor feedback on written work in an online program.

International Journal on Elearning, 7(2), 311-329.

Zeidner, M., & Matthews, G. (2005). Evaluation anxiety. In A. J. Elliot, & C. S. Dweck (Eds.),

Handbook of competence and motivation (pp. 141–163). London: Guildford Press.

Zhu, C. (2012). Student satisfaction, performance, and knowledge construction in online

collaborative learning. *Educational Technology & Society, 15*(1), 127-136.

Zimmerman, B.J. (2001). Theories of self-regulated learning and academic achievement: An

overview and analysis. In *self-regulated learning and academic achievement: theoretical*

perspectives, 2nd ed, ed. B.J. Zimmerman and D.H. Schunk, 1-36. Mahwah, NJ.

APPENDIX A

CONSENT LETTERS

CONSENT TO PARTICIPATE IN PRE-SURVEY

Title of Study**: Exploring Students' Experiences with mTuner**

You are asked to participate in a research study conducted by Lindsay Shaw and Dr. Jill A. Singleton-Jackson, from the University of Windsor as a fulfilment of a Master's thesis.

If you have any questions or concerns about the research, please feel to contact Lindsay Shaw at shaw11r@uwindsor.ca.

PURPOSE OF THE STUDY
The objective of the study is to find out about students' experiences with mTuner, the online assessment tool being implemented in introductory to psychology courses. More specifically, we hope to discover students' attitudes toward the tool and whether it is beneficial to the learning process.

PROCEDURES
If you volunteer to participate in this study, you will be asked to complete a pre-test survey prior to taking any mTuner midterms that will take no more than 30 minutes. You will also have the opportunity in the future to participate in a post-survey and focus group. Given the mTuner implementation, only those enrolled in the introductory to psychology course will be able to participate. Dr. Singleton-Jackson will be a co-researcher in the project, but will not have access to any identifying information. Therefore you will remain anonymous from her, and your participation in this study will in no way impact your final grade.

POTENTIAL RISKS AND DISCOMFORTS
The online surveys will ask you questions regarding everyday academic behaviour, which poses minimal risk. Your instructor of the course will be a co-researcher on this project, which risks a dual-relationship. However, Dr.Singleton-Jackson will not have access to any of the data until it has been anonymized, and any identifying information is deleted. With that, it is important to note that your participation in this study will have no impact on your performance in the course, neither positively or negatively.

POTENTIAL BENEFITS TO PARTICIPANTS AND/OR TO SOCIETY

First and foremost you will have the opportunity to express your opinion on assessment, something that will continue to influence your academics. The findings from this study will be sent to the developer of mTuner, which may help in their improvement of the tool. In other words you will not only have a hand in research, but practical educational development, a very rewarding opportunity. Further, there are several other research labs in the psychology department working to explore similar topics. The findings of this study may be influential in the development of interventions and/or surveys.

COMPENSATION FOR PARTICIPATION

Participants will receive 0.5 bonus points for 30 minutes of participation towards the psychology participant pool, if registered in the pool and enrolled in one or more eligible courses
You have the opportunity to participate in future phases of the study in which you will receive more bonus points. You will be notified of these phases.

CONFIDENTIALITY

Any information that is obtained in connection with this study and that can be identified with you will remain confidential and will be disclosed only with your permission.

. You will need to provide your name so that participant points will be awarded. By consenting to the study, you will also be consenting to the release of your grades. Once the grades have been submitted and linked to your data, your names, and any other identifying information will be deleted. Dr. Singleton-Jackson will not have access to the data until your identifying information has been deleted. You will use self-generated identification codes for both the online surveys so you remain anonymous from Dr.Singleton-Jackson, and in turn your information will be confidential. With that, the information you provide will in no way impact your performance in the class.

PARTICIPATION AND WITHDRAWAL

During the online survey you are free to skip any questions you don't wish to answer. You may also withdraw your participation at any time during the study by just exiting your browser. By withdrawing from the study, you will not receive any bonus points as you must reach the end of the survey to provide identifying information to receive the credit. If you complete less than 80% of the questionnaire or engage in non-serious responding you will only receive partial credit. If you wish to withdrawal from the study after completion and submission of the online surveys you must notify the lead researcher within 3 days of the submission.

FEEDBACK OF THE RESULTS OF THIS STUDY TO THE PARTICIPANTS

Research findings will be made available to you. The findings will be posted on the REB website. As no contact information will be collected for you, you will have to go to the site and look at the findings should you choose to do so, but the researcher cannot contact you to inform you that the findings have been posted.

Web address: www.uwindsor.ca/reb

Date when results are available: no later than September 1, 2017

SUBSEQUENT USE OF DATA

These data may be used in subsequent studies, in publications and in presentations.

RIGHTS OF RESEARCH PARTICIPANTS

If you have questions regarding your rights as a research participant, contact: Research Ethics Coordinator, University of Windsor, Windsor, Ontario, N9B 3P4; Telephone: 519-253-3000, ext. 3948; e-mail: ethics@uwindsor.ca

CONSENT TO PARTICIPATE
Please indicate your willingness to participate in the study:

- ☐   I agree to participate in the study
- ☐   I do not wish to participate in the study

- ☐   Please check the box if you understand the information provided for the study **Exploring Students' Experiences with mTuner** as described herein. My questions have been answered to my satisfaction, and I agree to participate in this study. I have been given a copy of this form.

SIGNATURE OF INVESTIGATOR

These are the terms under which I will conduct research.

_____   _____
Signature of Investigator   Date

University of Windsor

CONSENT TO PARTICIPATE IN POST-SURVEY

Title of Study**: Exploring Students' Experiences with mTuner**

You are asked to participate in a research study conducted by Lindsay Shaw and Dr. Jill A. Singleton-Jackson, from the University of Windsor as a fulfilment of a Master's thesis.

If you have any questions or concerns about the research, please feel to contact Lindsay Shaw at shaw11r@uwindsor.ca.

PURPOSE OF THE STUDY
The objective of the study is to find out about students' experiences with mTuner, the online assessment tool being implemented in introductory to psychology courses. More specifically, we hope to discover students' attitudes toward the tool and whether it is beneficial to the learning process.

PROCEDURES
If you volunteer to participate in this study, you will be asked to complete a post-test survey that will take no more than 30 minutes. You will also have the opportunity in the future to participate in a focus group. Given the mTuner implementation, only those enrolled in the introductory to psychology course will be able to participate. Dr. Singleton-Jackson will be a co-researcher in the project, but will not have access to any identifying information. Therefore you will remain anonymous from her, and your participation in this study will in no way impact your final grade.

POTENTIAL RISKS AND DISCOMFORTS
The online surveys will ask you questions regarding everyday academic behaviour, which poses minimal risk. Your instructor of the course will be a co-researcher on this project, which risks a dual-relationship. However, Dr.Singleton-Jackson will not have access to any of the data until it has been anonymized, and any identifying information is deleted. With that, it is important to note that your participation in this study will have no impact on your performance in the course, neither positively or negatively.

POTENTIAL BENEFITS TO PARTICIPANTS AND/OR TO SOCIETY

First and foremost you will have the opportunity to express your opinion on assessment, something that will continue to influence your academics. The findings from this study will be sent to the developer of mTuner, which may help in their improvement of the tool. In other words you will not only have a hand in research, but practical educational development, a very rewarding opportunity. Further, there are several other research labs in the psychology department working to explore similar topics. The findings of this study may be influential in the development of interventions and/or surveys.

COMPENSATION FOR PARTICIPATION

Participants will receive 0.5 bonus points for 30 minutes of participation towards the psychology participant pool, if registered in the pool and enrolled in one or more eligible courses
You have the opportunity to participate in future phases of the study in which you will receive more bonus points. You will be notified of these phases.

CONFIDENTIALITY

Any information that is obtained in connection with this study and that can be identified with you will remain confidential and will be disclosed only with your permission.

You will need to provide your name so that participant points will be awarded. By consenting to the study, you will also be consenting to the release of your grades. Once the grades have been submitted and linked to your data, your names, and any other identifying information will be deleted. Dr. Singleton-Jackson will not have access to the data until your identifying information has been deleted.. You will use self-generated identification codes for both the online surveys so you remain anonymous, from Dr.Singleton-Jackson and in turn, your information will be confidential. With that, the information you provide will in no way impact your performance in the class.

PARTICIPATION AND WITHDRAWAL

During the online survey you are free to skip any questions you don't wish to answer. You may also withdraw your participation at any time during the study by just exiting your browser. By withdrawing from the study, you will not receive any bonus points as you must reach the end of the survey to provide identifying information to receive the credit. If you complete less than 80% of the questionnaire or engage in non-serious responding you will only receive partial credit. If you wish to withdrawal from the study after completion and submission of the online surveys you must notify the lead researcher within 3 days of the submission.

FEEDBACK OF THE RESULTS OF THIS STUDY TO THE PARTICIPANTS

Research findings will be made available to you. The findings will be posted on the REB website. As no contact information will be collected for you, you will have to go to the site and look at the findings should you choose to do so, but the researcher cannot contact you to inform you that the findings have been posted.

Web address: www.uwindsor.ca/reb

Date when results are available: no later than September 1, 2017

SUBSEQUENT USE OF DATA

These data may be used in subsequent studies, in publications and in presentations.

RIGHTS OF RESEARCH PARTICIPANTS

If you have questions regarding your rights as a research participant, contact: Research Ethics Coordinator, University of Windsor, Windsor, Ontario, N9B 3P4; Telephone: 519-253-3000, ext. 3948; e-mail: ethics@uwindsor.ca

CONSENT TO PARTICIPATE
Please indicate your willingness to participate in the study:

☐ I agree to participate in the study
☐ I do not wish to participate in the study

☐ Please check the box if you understand the information provided for the study **Exploring Students' Experiences with mTuner** as described herein. My questions have been answered to my satisfaction, and I agree to participate in this study. I have been given a copy of this form.

SIGNATURE OF INVESTIGATOR

These are the terms under which I will conduct research.

_____   _____
Signature of Investigator   Date

APPENDIX B

LEARNING-AND GRADE-ORIENTATION (LOGO II) SCALE

Part I: 5-point Likert-type scale (1=strongly disagree to 5=agree strongly)

*bolded items indicate learning-oriented features

1. **I enjoy classes in which the instructor attempts to relate material to concerns beyond the classroom.**
2. I think it is unfair to test students on material not covered in class lectures and discussions, even if it is in reading assignments.
3. I dislike courses which require ungraded out-of-class activities.
4. **I prefer to write a term paper on interesting material then to take a test on the same general topic.**
5. **I get annoyed when lectures or class presentations are only rehashes of easy reading assignments.**
6. Written assignments (i.e., homework, projects, etc.) that are not graded are a waste of a student's time.
7. I **appreciate the instructor who provides honest and detailed evaluation of my work though such evaluation is sometimes unpleasant.**
8. I think that without regularly scheduled exams I would not learn and remember very much.
9. Instructors expect too much out-of-class reading and study by students.
10. **I find the process of learning new material fun.**
11. I dislike courses in which a lot of material is presented in class, or in readings, that does not appear on exams.
12. **Easy classes that are not pertinent to my educational goals generally bore me.**
13. **A teacher's comments on an easy test mean more to me than my actual test scores.**
14. I do not find study at home to be interesting or pleasant.
15. **I am more concerned about seeing which questions I missed than I am with finding out my test grade.**
16. I think grades provide me a good goal to work toward.

Part II: 5-point Likert-type scale (1=never to 5=always)

17. **I do optional reading that my instructors suggest even though I know it won't affect my grade.**
18. **I try to make time for outside reading despite the demands of my coursework.**
19. I try to get old tests when I think the instructor will use the same questions again.
20. I will withdraw from an interesting class rather than risk getting a poor grade.
21. I get irritated by students who ask questions that go beyond what we need to know for exams.

22. I stay after interesting classes to discuss material with the instructors.

23. I discuss interesting material that I've learned in class with my friends or family.

24. When looking at a syllabus on the first day of class, I turn to the section on tests and grades first.

25. I participate in out-of-class activities even when extra-credit is not given.

26. I buy books for courses other than those I am actually taking.

27. I borrow old term papers or speeches from my friends to meet class requirements.

28. I cut classes when confident that lecture material will not be on an exam.

29. I try to keep all my old textbooks because I like going back through them after class is over.

30. I try to find out how easy or hard an instructor grades before signing up for a course.

31. I'm tempted to cheat on exams when I'm confident I won't get caught.

32. I browse in the library when not working on a specific assignment.

APPENDIX C

ACADEMIC ENTITLEMENT QUESTIONNAIRE (AEQ)
(1=strongly disagree to 7=strongly agree).

1. If I don't do well on a test, the professor should make tests easier or curve grades.
2. Professors should only lecture on material covered in the textbook and assigned readings.
3. If I am struggling in a class, the professor should approach me and offer to help
4. It is the professor's responsibility to make it easy for me to succeed.
5. If I cannot learn the material for a class from lecture alone, then it is the professor's fault when I fail the test.
6. I am a product of my environment. Therefore, if I do poorly in class, it is not my fault.
7. I should be given the opportunity to make up a test, regardless of the reason for the absence.
8. Because I pay tuition, I deserve passing grades.

APPENDIX D

COGNITIVE TEST ANXIETY SCALE (CTA)

(Cassady et al., 2002)

(Not at all typical of me (score=1), Only somewhat typical of me (score=2), Quite typical of me (score=3), and Very typical of me (score=4)

1. I lose sleep over worrying about examinations.

2. While taking an important examination, I find myself wondering whether the other students are doing better than I am.

3. I have less difficulty than the average college student in getting test instructions straight. *

4. I tend to freeze up on things like intelligence tests and final exams.

5. I am less nervous about tests than the average college student. *

6. During tests, I find myself thinking of the consequences of failing.

7. At the beginning of a test, I am so nervous that I often can't think straight.

8. The prospect of taking a test in one of my courses would not cause me to worry.*

9. I am more calm in test situations than the average college student.*

10. I have less difficulty than the average college student in learning assigned chapters in textbooks. *

11. My mind goes blank when I am pressured for an answer on a test.

12. During tests, the thought frequently occurs to me that I may not be too bright.

13. I do well in speed tests in which there are time limits. *

14. During a course examination, I get so nervous that I forget facts I really know.

15. After taking a test, I feel I could have done better than I actually did.

16. I worry more about doing well on tests than I should.

17. Before taking a test, I feel confident and relaxed. *

18. While taking a test, I feel confident and relaxed. *

19. During tests, I have the feeling that I am not doing well.

20. When I take a test that is difficult, I feel defeated before I even start.

21. Finding unexpected questions on a test causes me to feel challenged rather than panicky. *

22. I am a poor test taker in the sense that my performance on a test does not show how much I really know about a topic.

23. I am not good at taking tests.

24. When I first get my copy of a test, it takes me a while to calm down to the point where I can begin to think straight.

25. I feel under a lot of pressure to get good grades on tests.

26. I do not perform well on tests.

27. When I take a test, my nervousness causes me to make careless errors.

*Items should be recoded to produce consistency in scale so that high values always reflect high-cognitive test anxiety responses. The possible range of scores is 27-108. Higher scores indicate higher cognitive test anxiety.

APPENDIX E

MODIFIED STUDENTS' CONCEPTION OF ASSESSMENT SCALE (Version VI)
(Brown et al., 2009)

*Please note traditional multiple choice assessments are defined as multiple choice assessments using paper and pencil.
For the mTuner version of this scale all "traditional multiple choice assessments" will be replaced with "mTuner".

Affect/Social Benefit (1=strongly disagree to 6=strongly agree)

1. Traditional multiple choice assessment is an enjoyable experience for me
2. Traditional multiple choice assessment is beneficial for me
3. Traditional multiple choice assessment makes me more motivated to learn

Irrelevance (1=strongly disagree to 6=strongly agree)

4. Traditional multiple choice assessments have little impact on my learning
5. Traditional multiple choice assessments are unfair to students
6. I ignore my Traditional multiple choice assessment results

Improvement (1=strongly disagree to 6=strongly agree)

7. My teachers use Traditional multiple choice assessment to help me improve
8. I look at what I got wrong or did poorly on to guide what I should learn next after Traditional multiple choice assessments
9. I make use of the feedback I get to improve my learning after Traditional multiple choice assessment

External Attributions (1=strongly disagree to 6=strongly agree)

10. Traditional multiple choice assessment is a way to determine how much I have learned from teaching
11. Traditional multiple choice assessment predicts my future performance
12. Traditional multiple choice assessment results show how intelligent I am

mTuner version only*

13. Which feature of mTuner did you find most helpful? (1=having hints after getting a question wrong, 2=being able to answer a question a second time for fewer marks, 3=receiving immediate feedback after each question, 4=being able to recall as much information as possible at the beginning of each question, 5=none of the features were helpful)

14. Which assessment do you prefer more? (1=traditional, 2=I like them equally, 3=mTuner)
15. In an open ended response, please describe in further detail your preference (i.e. I prefer traditional assessment over mTuner because I do not feel comfortable using computers).

APPENDIX F

ENGAGEMENT SALE

(1=never, 2=almost never, 3=sometimes, 4=almost every time, 5= everytime)

1. During the current mTuner midterm how often did you
   a. Read the feedback given after each multiple choice question
   b. Read/listen to the e-text or video lecture that assisted you in answering the question.
   c. Answer the question a second time after getting the question wrong
   d. Use the open textbox to write after each multiple choice question
2. Prior to the mTuner assessment how often did you
   a. Seek advice from professor/graduate assistant on how to use mTuner
   b. Ask other students to help you understand mTuner
   c. Help peers on how to use mTuner
   d. Work with peers during the mTuner midterm.

APPENDIX G

MTUNER INSTRUCTION SHEET

mTuner Instructions

PLEASE take the time to <u>read ALL of this document</u> so that you can avoid stress for yourself and for us!

* mTuner #3 will not be available until 5PM on Thursday, March 23 Do not try to log in prior to this time and date!!!

<u>REMEMBER:</u> The goal here is for you to understand the material, to strengthen your memory connections, so that you are gaining knowledge and LEARNING. As a result, you will have a better chance of passing the cumulative final, and therefore making you a happy student.

In order to successfully complete your mTuner assignment you must:

1) Use a computer – mTuner will not work with mobile devices and tablets
2) **USE FIREFOX** – mTuner is not compatible with other browsers

3) **Have popups enabled in your browser** – there will be hints that pop up if you answer incorrectly
4) Have an update Adobe Flash Player to be able to see your hints

5) Have the Textbook in front of you OR be logged into your e-text prior to beginning your assignment

   If you are unsure if your computer has these requirements, please use a computer in Leddy Library. All library computers are equipped with Firefox and Adobe Flash Player; still make sure that popups are enabled on the computer you are using—do not hesitate to ask for help from the library staff if you are uncertain about how to do something

Step-by-step Instructions

Step 1

- Go to uwindsor.mtuner.ca

Step 2
- Log in using your email ID and student number
- **Use ONLY the ID, <u>do not use the @uwindsor.ca part</u>**
- Example: My email address is jjackson@uwindsor.ca and my student number is 100123456
  - ID: jjackson   Password: 100123456

  - *<u>DO NOT USE THE @UWINDSOR part of your email address!!!!!!</u>*

Step 3

- Select "Intro to Psych as a Social Science (46-116-**01**)" OR "Intro to Psych as a Social Science (46-116-**02**)"
- Section 01 is the T/TH class
- Section 02 is the M/W class

- You will see "mTuner Assignment 3 – Chapters 8 and 14" and will be able to click start once the time slot opens up.

Note

This mTuner assignment will be open for 48 hours—from Thursday, March 23 at 5:00

PM until Saturday, March 25 until 5:00 pm.
Once you begin the assignment you will have **3 *hours*** to complete it.

It is expected that once you begin the assignment you will finish it in one sitting. Each question contains a primer, a first chance at answering the multiple choice

question, a hint if you answer incorrectly on your first attempt, a second chance to answer the multiple choice question, and finally, an explanation.

- You are not graded on how you answer the primer question, this is only intended to get you thinking about the concepts and ideas involved in the question

- For those who prefer the hardcopy of the text as opposed to the e-text: When a hint pops up that links you to the e-text, simply close the pop-up window and review the information in the Hint box – The page number will be written there

  for you to investigate in your hardcopy of the text.

- o Each question is worth 2 marks; therefore, if you answer the question correctly on your first attempt you will receive 2 points. If you answer incorrectly on your first attempt you will be given a hint and a second attempt to answer the question; if you then get the answer correct you will receive ½ credit. If you answer incorrectly on your second attempt you will receive 0 points on that question.

Each stage in answering a question is timed so be mindful; however, after each question explanation you are prompted to a screen that asks you to **"quit" or "continue".** You should select continue each time, but you can also use this break to go to the washroom

or grab a drink before selecting "continue".

Some of the primer questions may not be open-ended; still, use this time to consider the main concepts or ideas being discussed in the question so as to prepare yourself to

answer the multiple choice when you move on to your first attempt.

Be aware that all of your activity on the mTuner platform is recorded and visible to us on the back end. As such, it is highly recommended that you uphold your integrity as a student and not engage in fraudulent conduct. Any suspicious activity will be noted and investigated. Having said that, there should be no need to try to scam your way through this chance you are being given. You are provided with ample opportunities to succeed so it would be foolish to act against that.

If you are experiencing major difficulties, please contact [the instructor] or [the head GA].

Good Luck

VITA AUCTORIS

Lindsay Shaw was born in 1993 in Oshawa, Ontario. She graduated from Holy Names High School in 2011. From there she attended Wilfrid Laurier University where she obtained an Honours Bachelor of Arts degree in 2015, majoring in Psychology: Research Specialist Stream, with a minor in Women and Gender Studies. She is currently a Master's candidate in the Applied Social Psychology program at the University of Windsor and expects to graduate in the Fall of 2017.