# FIRST STEPS TOWARDS IMPLEMENTING UNIVERSAL DESIGN FOR LEARNING TO SUPPORT EQUITABLE ASSESSMENTS

Nikolas Zelem, Diana Skrzydlo and Chelsea Uggenti
University of Waterloo, Canada
cuggenti@uwaterloo.ca

*Universal Design for Learning (UDL) promotes inclusion of a diverse set of student learning needs and is beneficial for improving student learning outcomes regardless of physical or neurological ability. Yet instructors may ask themselves, "Where do I start?" in terms of implementing UDL strategies in their courses. A review of relevant literature for the application of UDL strategies for assessments in post-secondary mathematical and statistical education is provided. A list of nine basic changes made by instructors to improve the accessibility and inclusivity of assessments in their courses is offered. Such changes are intended to provide immediate impact with relatively low effort aimed towards instructors with minimal UDL experience. Two case studies focusing on the implementation of UDL strategies for assessments in statistics courses are included for reference. This paper serves as an introduction into the realm of UDL and, more specifically, UDL practices for assessments.*

## INTRODUCTION

Good teaching is inclusive teaching. As post-secondary mathematics and statistics instructors seek to provide education to a wide variety of students, it is important to keep equity and accessibility in mind. This paper intends to collect UDL practices that instructors can easily incorporate into their assessments. In this literature review, several guiding principles for implementing UDL in assessments are discussed. Student attitudes towards UDL practices are also explored, informing the subsequent list of nine recommended strategies for instructors to start with when implementing UDL practices in their assessments. The strategies are listed, following the literature review, in ascending order of required instructor efforts. To provide a tangible example of how these strategies look in practice, two case studies focusing on the implementation and impact of UDL for assessments in statistics courses offered by the Department of Statistics and Actuarial Science at the University of Waterloo (a large post-secondary institution in Canada) are provided.

## LITERATURE REVIEW

The purpose of this literature review is to compile a list of strategies – focusing on UDL practice – that instructors can use to improve the accessibility of their assessments for all students. This section provides an overall description of the state of literature surrounding UDL in post-secondary education. While our review was not systematic, some important broad observations are shared.

First, there exists a considerable amount of material concerning UDL in primary and secondary education, with less information available for post-secondary education. UDL in post-secondary STEM education has been covered in some detail, but statistics-specific papers are exceedingly uncommon.

The terminology surrounding universal design in post-secondary education is somewhat inconsistent. Several comprehensive literature reviews of papers involving UDL have also found inconsistencies with the use of terminology. Rao et al. (2014) note that while the phrase "universal design for learning" commonly refers to the model described by CAST (2018), it has also been used as an umbrella term for the application of other universal design frameworks. Two examples are Universal Design of Instruction (UDI) and Universal Instructional Design (UID). McGuire et al. made a similar observation in 2006, going so far as to describe the large variety of acronyms as an "alphabet soup" (p. 172) and asserting that anyone researching or practicing universal design in an educational context should be "intentional in their use of terminology" (p. 172), a statement worth echoing in 2023.

Finally, the content of many UDL papers is purely descriptive in nature; empirical data concerning the efficacy of UDL practices is rare – even more so in a post-secondary STEM context. A large proportion of the literature has thus far been dedicated to either simply describing UDL practices or examining students' attitudes toward them or perceptions of them. The 2014 literature review by Rao et al. found only 13 articles that measured the effect of a universal design intervention. Of those, less than two-thirds used quantitative methods. While this field may still be in its infancy, quantitative

research examining whether UDL practices have any effect on student performance and wellbeing will need to be carried out for progress to be made.

The topic of Universal Design of Assessment (UDA) and how it differs from UDL is subsequently considered. According to the American Educational Research Association [AERA], the American Psychological Association (APA), and the National Council on Measurement in Education [NCME] (2014, as cited in Lazarus et al., 2022), universally designed assessments "should improve the validity, reliability, and fairness of assessments for all students" (p. 3). This definition suggests that UDA applies to the technical aspect of assessments rather than the supports that surround assessments. Additionally, many of the guidelines for UDA found in reports such as the one from Lazarus et al. are intended specifically for summative assessments. Many of the suggestions made in this paper deal with the supports that surround assessments – as opposed to the assessments themselves – and a variety of assessment types. For these two reasons, the more general term, UDL, will be used throughout the remainder of this paper.

With an understanding of the landscape established, we examine the UDL literature starting with a paper that offers a general suggestion for implementing UDL strategies. Tobin (2019) mentions promoting a "plus one" (p. 16) mindset when applying UDL principles. While Tobin was referring to course content, the same approach can be applied to assessments – the goal being to improve assessments every time they are administered, even if the improvement is relatively small. A good place to start is identifying areas of student weakness. Tobin (2019) describes three "hot spots" (p. 16) that might demand special attention: sections of course material where students tend to (1) have many questions; (2) make mistakes on assessments; and (3) need ideas explained in multiple ways. Focusing on UDL practices that address any number of these areas can help to maximize return on investment.

One of the few papers that examined UDL practices in undergraduate statistics was Block's 2012 study, which investigated the effects of open-book testing on student performance and attitudes. The study involved approximately 260 students for each of the four semesters over which it was performed. Open-book tests were not well-liked; students resented the fact that the material on open-book tests was more challenging. However, offering small supports, such as allowing students to bring a hand-written notecard, resulted in improved student performance and higher levels of enjoyment compared to completely open-book tests. The purpose of a small memory aid such as a notecard is to reduce the stress associated with unnecessary memorization and prevent students from having to expend excessive cognitive energy on recall. Provided that the goal of an exam is not to test basic knowledge, notecards or other small memory aids are one of the best extra resources that can be easily offered by an instructor.

Another avenue for improving accessibility on tests and exams involves the time allotted to complete them. Some students might not be able to process information at the same speed as others and may take longer to complete tests as a result. Ketterlin-Geller & Ellis (2020) suggest that more time can be provided to those students to increase the accessibility of the assessment. However, providing supports only to students who request them requires extra effort on the part of the instructor, and may still leave some students behind. Instead, it may be worth increasing the time given to all students or reducing the length of the assessment.

Support for a test can also be offered before the test takes place. In 2014, Kumar & Wideman conducted a study intended to determine student perceptions of UDL interventions on course accessibility. The study focused on 50 students in an undergraduate health sciences course, but the results could still potentially be generalized. They found that 97% of students accessed study guides that included topic outlines and lists of key concepts. Of that 97%, 94% found topic outlines helpful and 100% found the lists of key concepts helpful. One student said that the study guide helped her to "narrow [her] focus" (p. 136), which helped to reduce her stress. The same student also said "sometimes students are not really motivated to learn if they're not sure what the expectations are" (p. 137). These results indicate that offering a study guide for major assessments like midterm and final exams could benefit most students – regardless of ability. Study guides do not require a large amount of time to produce, making them an easy addition to a course.

Study guides function in a similar way to self-assessment questions in that they make students aware of what material is covered by a test, although self-assessment questions do not necessarily have to be associated with a test. A study conducted by Nieminen & Pesonen in 2020 involved interviewing three undergraduate STEM students with disabilities that took a mathematics class that utilized many

UDL strategies. All three students identified self-assessment practices as being beneficial to their learning.

The three students in Nieminen & Pesonen's 2020 study also found that a "course rubric with exemplars" provided at the beginning of the course helped them to set goals and ensure they were understanding each topic. Black et al. (2015) also conducted interviews – though their sample of 15 students included those with and without disabilities – in which students mentioned that in order for them to achieve at their expected level, they need instructors to clearly define course objectives, otherwise time management would become more difficult. These findings suggest that defining clear expectations for a course is valuable.

It is also important to define expectations at a higher level of detail. Smith (2012) found that 80 graduate students in a research methods course tended to prefer to be able to refer to an exemplar when completing an assignment and/or have access to a rubric or template outlining what is expected of them. These same students also prefer to receive constructive feedback on a frequent basis. While giving feedback on assignments does require a significant time commitment from instructional staff, rubrics and/or templates are not difficult to create, and exemplars can be easily collected from students. These small changes can help support students on assignments by giving them an idea of what they should be doing.

While many of the papers found in this literature review suggest changes related to tests and exams, it is also important to consider other assessment types. Projects, for example, are common in undergraduate statistics courses because they offer a chance for students to work with data. Dahlstrom-Hakki & Wallace (2022) investigated strategies to help students with learning disabilities, ADHD, and autism in statistics courses. Their study initially involved 68 students, but only 39 were included in the published results. The strategies they implemented use all three suggestions from CAST (2018), but the most valuable takeaway focused primarily on engagement. Instructors prepared datasets that the students were interested in or datasets that were comprised of data that was collected from or by the students. This approach can give students a sense of agency that may be missing otherwise.

Another way to engage students is to have them take control of their learning. Interviews conducted by Black et al. (2015) found that students placed considerable value on the freedom to choose how they would express their learning, believing that this freedom to choose the method of expression helps them to demonstrate their learning effectively. While replacing a test with a project takes a considerable amount of preparation time for the instructor, offering students their own choice of medium within an assessment type (e.g., oral presentation vs. recorded video) does not require a large time investment on the part of the instructor.

Kumar & Wideman also examined the effects of increased student choice for projects in their 2014 study involving students in an undergraduate health science course. They gave students their choice of topic on projects, receiving a widely positive response. Students were also given the choice of due dates, the presentation format, and, for some projects, the choice to collaborate or work independently – all of which saw the same positive response as the choice of topic. Student comments were also positive with many students remarking on the flexibility that these choices offered them. One student pointed out that with this flexible learning approach the student can choose "the option that would be best for [their] learning" (Kumar & Wideman, 2014, p. 135).

Another form of flexibility comes from course design. Optional assignments were included as a part of the course in Kumar & Wideman's 2014 study. These short assignments gave students the opportunity to reduce the weight of the final exam by completing them, adding some flexibility to the grading scheme. There were four assignments in total, with 43 (86%) students completing all four and 6 (12%) completing three. Almost all students were eager to reduce the weight of the final exam, suggesting that students respond positively to more flexible grading schemes. Interviews with a small subset of students taking part in Kumar & Wideman's study revealed that students believed that the UDL practices contributed to their perceived success in the course.

LIST OF BASIC UDL CHANGES

Informed by the literature review, we created a list of nine adjustments that instructors teaching post-secondary mathematics and statistics courses can make to improve the accessibility and inclusivity of their course assessments for all students. These changes are listed from little-to-no instructor burden up to more instructor burden. Both the results from the literature review and previous instructor

experience were utilized in its development. We offer a further explanation and example(s) after each change. It is important to note that our list is by no means exhaustive and that other strategies may improve the universal design of course assessments.

- *Group Work/Collaboration Options*
  Provide options for students to work in pairs/small groups to submit assignment/project (or quiz/test). Reduces grading burden; encourages authentic collaboration.
- *Flexible Grading*
  Allow multiple attempts at online auto-graded assessments; drop lowest X of Y similar assessments automatically; adjust weight of midterm(s) based on performance. LMS assists by handling grade calculations and any key assessments that cannot be dropped can be specified.
- *Flexible Deadlines*
  Allow students some (limited) flexibility when submitting assessments. E.g., "slip days" or late assignment "tokens" (require tracking), automatic extensions, or extensions if student sought help.
- *Extra Resources on Tests*
  Provide memory aids or helpful technology to all students. E.g., reference or formula sheet (created by instructor/student), calculator, dictionary, or even a word processor/computer.
- *Extra Time on Tests*
  Provide additional time to lower time-related stress barriers for all students. If it is logistically difficult to book additional classroom time, design a shorter test given the time available.
- *Self-Assessment/Diagnostic Questions*
  Create and provide students with self-assessment questions to check their own knowledge preparedness. E.g., questions from past terms, common student misconceptions/mistakes, or a list of things they should know how to do. Answers may or may not be included (useful either way).
- *Clarity of Expectations*
  Give explicit instructions about how/when to submit any assessment. Link assessment to learning goals for the course (consider a "what/how/why"). When necessary, provide a rubric and examples of previous student work (with their consent) to show students what is expected of them. If any specific software is required, instructions or practice on how to use it are valuable.
- *Choice of Topic*
  Allow student agency in the choice of a meaningful topic for an assessment. Encourages students to engage authentically with course material. E.g., Provide a list of potential topics, allow any topic within specific parameters, or leave topic(s) completely open. The importance of providing clarity of expectations increases as flexibility/openness increases.
- *Choice of Deliverable Medium*
  Based on learning outcomes assessed by project/assignment, consider whether they could be demonstrated in multiple ways. E.g., if written communication is important, allow students to choose between a report, poster, or pamphlet. If oral communication is the goal, allow students to choose between an in-person presentation or pre-recorded video. Allowing multiple means of expression lets students demonstrate their knowledge in the way that works best for them.

CASE STUDIES

Implementing the previous list of basic UDL changes to assessments may seem daunting at first. Instructors may wonder: Should I incorporate all the changes listed? What do they look like in practice? The following two case studies provide tangible examples of how these strategies, in particular flexible grading, group work, self-assessment questions, and choice of project topic and deliverable, were implemented in different statistics courses. Their impacts are also mentioned.

In Winter 2023 for a large (250 students) second-year introductory statistics course for science students, online quizzes for knowledge checks were utilized. A total of 8 quizzes were set throughout the term, with only 5 counting towards the final grade. Students had ~10 days to complete a quiz, where 5 questions were asked in each quiz. They had unlimited attempts to complete the quiz before the

deadline with each quiz session lasting for 30 minutes (students could start a new 30-minute quiz session right after the previous one). Questions with correct answers were carried forward with every attempt so that students did not have to repeat those questions. Most questions contained hints/feedback that students could access during the quizzes to help them in real time. After a deadline passed, students had access to those quizzes, now not worth any marks, and could try multiple attempts at the questions in preparation for other assessments (e.g., in-person tests/midterms). In general, students noted how helpful these quizzes were to check their understanding of the material each week, since most weeks consisted of a quiz. They also found it helpful to have access to the quizzes after the deadlines to continue with their practice. No accommodations through the university's accessibility services were required for students due to the flexibility in the assessment structure.

In Fall 2022 for a medium (60 students) third-year probability models course for business students, the flipped classroom (Talbert, 2017) model was used. A series of "Knowledge Check" questions were provided with each week of content videos so students could test their own knowledge and preparation. Numerical answers (but not full solutions) were given so they could check their work. In the weekly in-person meeting, a short, ungraded quiz on the video content was administered using Kahoot! – a free game-based learning platform – for students to gauge their own preparedness and ask questions. During these meetings, weekly activities were completed in groups that remained the same for the term so students could form strong bonds with their group members. The activities varied from week to week, but always related to key learning outcomes and threshold concepts that students had struggled with in past terms. Groups handed in one page with everyone's name on it, and solutions were posted immediately after the meeting time. Because a busy in-person class may not be the best environment for every student every week, there was also the option to submit the tutorial activity online within 24 hours to get full credit. As well, the lowest two activities were dropped, so two could be missed without penalty. It is telling that even towards the end of the term, the attendance at these tutorials was over 80%.

For graded assessments, students had 3 written assignments which could be completed in pairs. These included both theoretical and coding questions, as well as two new types: reflective questions about their progress and goals, and a "probability models inventory" where students were asked to identify their preferred example for each concept in the course. There were two proctored, individual-work tests (a midterm and a non-cumulative final exam) with one hour of extra time and 2 pages of handwritten reference material allowed for all students. There was a large group project where students had their choice of topic to model with a Markov Chain, and their choice of a video or in-person presentation. This gave students a large amount of agency and they were able to demonstrate their knowledge in a way that worked best for them. They also had to view and evaluate other groups' presentations and reflect on what they learned from the project. Finally, an individual 15-minute oral exam was used. Students were asked conceptual questions such as defining, comparing, and evaluating models. Since this was an unfamiliar assessment type for most students, practice questions were posted, and time was given in the last class for students to practice in pairs. Overall, the student course evaluations were highly complementary, with many noting they felt like the instructor really wanted them to succeed. A few students who normally requested accommodations did not feel they needed them, and engaged with the course exactly as it was designed, since the barriers they often face were not present.

## DISCUSSION

Universal Design for Learning is a broad, yet crucial, teaching approach that promotes and encourages the inclusion of a diverse set of student learning needs. We attempt to clarify how UDL strategies may be implemented into assessments for post-secondary mathematics and statistics courses. Our literature review provides examples of several practices and/or ideas for such applications, culminating in our list of nine basic changes that instructors may consider in order to improve the accessibility of their courses. These suggestions are intended to provide immediate impact with relatively low-to-moderate effort and are aimed at instructors with minimal UDL experience who are looking for a starting point. The case studies further illustrate how such practices can be applied to course assessments and student reactions towards them.

This research is introductory in nature as we endeavor to scratch the surface of UDL and its relationship with the field of mathematical and statistical education. Our work should not be regarded

as comprehensive nor systematic, rather a first step in the right direction. We recognize that other unmentioned strategies may be effective, or perhaps more effective, in improving accessibility in course assessments. The purpose of our list is to encourage instructors to implement some of the changes (one step at a time) in their courses and to share any beneficial UDL practices in assessments with the broader community.

REFERENCES

Black, D. R., Weinberg, L. A. & Brodwin, M. G. (2015). Universal design for learning and instruction: Perspectives of students with disabilities in higher education. *Exceptionality Education International, 25*(2), 1-26.

Block, R. M. (2012). A discussion of the effect of open-book and closed-book exams on student achievement in an introductory statistics course. *Problems, Resources, and Issues in Mathematics Undergraduate Studies, 22*(3), 228-238.

CAST (2018). *Universal design for learning guidelines version 2.2*. Retrieved from http://udlguidelines.cast.org

Dahlstrom-Hakki, I. & Wallace, M. L. (2022). Teaching statistics to struggling students: Lessons learned from students with LD, ADHD, and autism. *Journal of Statistics and Data Science Education, 30*(2), 127-137.

Ketterlin-Geller, L. R. & Ellis, M. (2020). Designing accessible learning outcomes assessments through intentional test design. *Creative Education, 11*, 1201-1212.

Kumar, K. L. & Wideman, M. (2014). Accessible by design: Applying UDL principles in a first year undergraduate course. *Canadian Journal of Higher Education, 44*(1), 125-147.

Lazarus, S. S., Johnstone, C. J., Liu, K.K, Thurlow, M. L., Hinkle, A. R. & Burden, K. (2022). *An updated state guide to universally designed assessments* (NCEO Report 431). National Centre on Educational Outcomes.

McGuire, J. M., Scott, S. S. & Shaw, S. F. (2006). Universal design and its applications in educational environments. *Remedial and Special Education*, *27*(3), 166-175.

Nieminen, J. H. & Pesonen, H. V. (2020). Taking universal design back to its roots: Perspectives on accessibility and identity in undergraduate mathematics. *Education Sciences, 10*(1), 12.

Rao, K., Ok, M. W. & Bryant, B. R. (2014). A review of research on universal design educational models. *Remedial and Special Education*, *35*(3), 153-166.

Smith, F. G. (2012). Analyzing a college course that adheres to the universal design for learning (UDL) framework. *Journal of the Scholarship of Teaching and Learning, 12*(3), 31-61.

Talbert, R. (2017). *Flipped learning: A guide for higher education faculty*. Stylus Publishing, LLC.

Tobin, T.J. (2019). Reaching all learners through their phones and universal design for learning. *Journal of Adult Learning: Knowledge and Innovation*, *4*(1), 9-19.