



A Panel Data Analysis of Using Multiple Content Modalities during Adaptive Learning Activities

Catherine A. Manly¹

Received: 1 March 2022 / Accepted: 20 February 2024
© The Author(s) 2024

Abstract

This study investigates the effect of offering multiple means of representing content, one aspect of Universal Design for Learning (UDL). Students across the full range of the disability spectrum all too often struggle to achieve course success. UDL-based course design promises to help students with disabilities (whether disclosed or not) as well as help all students, but research is needed to verify UDL's specific benefits for student learning outcomes. This study aims to better understand the efficacy of representing course content using multiple means, which is one aspect of UDL pertaining to perception. Data were gathered from over 50 online courses in 14 subjects across the undergraduate curriculum taught with an adaptive learning system at a women's institution. A panel data analysis with almost 200,000 cases of student learning activities investigated the effect of representing content in multiple ways (i.e., text, video, audio, interactive, or mixed content presentation format). When students used multiple modalities, a positive effect was found on student learning measures of knowledge gained within the adaptive system. The results have implications for future UDL-related research, as well as faculty development and curricular design.

Keywords Universal design for learning · Content modality · Quasi-experimental · College students · Students with disabilities · Undergraduate learning

Introduction

Considering the great variability among learners, finding instructional strategies that work well across ability levels holds importance for educators (Meyer et al., 2014). In the wake of COVID-19, many students have experienced learning challenges, including students in a variety of underperforming groups at all levels of education (Dorn et al., 2020; Herold & Chen, 2021; Manly et al., 2021; Zhang et al., 2021). This has magnified the need to iden-

✉ Catherine A. Manly
c.manly@fdu.edu

¹ Sammartino School of Education, Fairleigh Dickinson University, Teaneck, NJ, USA

tify effective strategies for assisting learning acquisition in college (Office for Civil Rights, 2021). The present analysis probes one promising technique posited to support students across the full range of ability levels, including students both with and without identified disabilities (Tobin & Behling, 2018).¹ Specifically, considering the Universal Design for Learning (UDL) framework's guideline to provide alternative modalities in representing content (CAST, 2018), I investigate the effect of use of multiple modalities on formative learning activities throughout a variety of online college courses.

Educational researchers have surprisingly little empirical evidence showing how specific educational design practices facilitated by new technological capabilities, such as recommended by UDL (Burgstahler, 2015), translate to learning success and other college outcomes (Kimball et al., 2016; Mangiardi & Serenelli, 2013). I address this by investigating student learning success in online courses using a learning-analytics-infused delivery system, which facilitates implementing some aspects of UDL. This framework, following efforts to improve universal design in the built environment (Hamraie, 2017), arose from work to improve educational circumstances for people with disabilities and is grounded in cognitive science (Rose, 2001). UDL theorizes that students benefit from multiple means of engagement, multiple means of representation, and multiple means of action and expression in their studies (CAST, 2018).

Importantly, UDL suggests that having content available in different modalities will result in beneficial outcomes. I investigate the causal effect of using multiple content modalities (i.e., text, video, audio, interactive, or mixed content) on student learning outcomes for undergraduates at a women's institution that serves predominantly older, nontraditional age students. Combining data from adaptive learning with other campus systems, I aim to discover whether the multiple content presentation recommended by UDL benefits these students. I undertake this research because the efficacy of using multiple content modalities as proposed by UDL still needs rigorous, empirical investigation in practice (Rao et al., 2014; Roberts et al., 2011). Development of a small research base has begun, such as Hall et al.'s (2015) work involving formative assessment in 14 middle school classes including a control group. However, that work looked more broadly at online versus offline reading. Thus, while some UDL research has addressed content representation, overall, few researchers have studied the connection between content presentation modalities and student learning outcomes, even as part of more comprehensive UDL research (Capp, 2017; Cumming & Rose, 2021; Rao et al., 2014). My research addresses this gap.

I conduct a theoretically informed panel data analysis of an authentic learning situation (Mayer, 2008), taking a short longitudinal approach. I look at change within student across two consecutive time points logged approximately 20 min or less apart across a single activity. I average across activities to investigate an overall effect posited across learning activities in a variety of courses. My goal is to identify the extent any beneficial effect exists when a student utilizes more than one modality when learning course content. My research question is: What are the effects of choosing more than one modality (either text, video, audio, interactive, or mixed) for learning course material on knowledge gain? I hypothesize that

¹ I respectfully acknowledge differing opinions about the relative merits of person-first or identity-first language when discussing people who have disabilities (Association on Higher Education and Disability, 2021). I use person-first language as a group signifier here since I do not discuss specific individuals who identify as having a disability.

use of more than one modality when learning content will have a substantively important positive effect on learning gain.

Theory and Literature Review

I frame my study by considering how online courses could be designed to support a universal range of abilities while simultaneously viewing each student individually and holistically. I draw upon the theoretical framework offered by UDL (Burgstahler & Cory, 2008), which facilitates systematic investigation and critique of intentional integration of different modalities in course design.

The concept of universal design came from the idea that designing a built environment accessible to people of all ability levels would produce a setting enabling rather than disabling participation by all individuals. In the early 1970s, architect Ron Mace attended North Carolina State University, where “Mace had to be carried up and down stairs to attend classes and was unable to use the men’s restroom because his wheelchair was too wide to fit through the entrance” (Evans et al., 2017, p. 277). This led him to pioneer the idea of barrier-free design, which he and others later expanded to the idea of universal design applying to everyone. Ramp structures like curb cuts offer a standard example of universal design, making it possible for people requiring wheelchairs to access spaces otherwise inaccessible via curbs or stairs. Although curb cuts were initially designed with disability access in mind, they are usable by and useful for many other non-disabled individuals, such as those wheeling a stroller, grocery cart, luggage, or hand truck.

Similarly, in higher education, advocates see universal design poised to become “a mainstream concern and a discourse serving the needs of students at large,” partly because “the wider objective of increasing diversity on campus is exceptionally well served by the model,” a conclusion drawn from an institutional case study of faculty, administrators, and other employees (Fovet & Mole, 2013, p. 124). Educational frameworks based on the idea of designing for a universal audience include Universal Design for Instruction (UDI), Universal Instructional Design (UID), and UDL (CAST, 2014; McGuire, 2014; Scott et al., 2003; Silver et al., 1998). In these frames, all people are viewed as having potential to benefit from design providing essential access to an otherwise disadvantaged subpopulation. This situates disability in the environment rather than individuals, whatever their current physical or mental capability and educational background (Evans et al., 2017). These frames also value providing content to students in multiple modalities. While parallels exist between these educational universal design formulations and each have different strengths, I focus on UDL because of its explicit articulation of multiple means of representation as a guiding principle, which can be explored empirically through offering options for perception as per the UDL guidelines.

UDL draws specifically upon brain imaging research, guiding learning design to facilitate academic achievement by diverse students whose capacities may vary significantly across the brain’s affective, recognition, and strategic networks (Rose, 2001). Recognizing that each student possesses a unique combination of strengths and weaknesses in each cognitive area enriches understanding of the dimensions across which human capability varies, informing design of educational experiences (Rose et al., 2006). However, UDL literature has more frequently focused on arguments for UDL’s importance than on empirical study

of its effects and effectiveness (Mangiatori & Serenelli, 2013; Roberts et al., 2011). What empirical research exists about universal design has often focused more on K-12 than college (Crevecoeur et al., 2014; Rao et al., 2014), and on perceptions or implementation activity rather than learning outcomes (Abell et al., 2011; Korterling et al., 2008; Lombardi et al., 2011). The modality use studied here falls under UDL's principle of providing multiple means of representation (CAST, 2014).² Within this, UDL recommends providing options for perception, connected to the brain's recognition capacity.

Modality Representation and UDL

UDL assumes that students enter college with a wide range of ability and prior experience and that students benefit from flexible paths to facilitate their learning. Courses designed with UDL in mind offer students multiple means of representation, including alternatives for auditory and visual information (CAST, 2014). The point of such multiplicity is not to offer additional complexity or add detrimental cognitive load (Beacham & Alty, 2006; Greer et al., 2013). Instead, UDL designers presenting options simultaneously should avoid unnecessarily increasing cognitive load which may otherwise increase barriers for some students (Kohler & Balduzzi, 2021). A design goal would be to allow students to pursue alternate paths through a course's content if they struggle to learn along the initial path or are functionally unable to follow a particular path.

This flexibility aligns well with availability of multiple modalities for alternative content presentation in adaptive systems such as in the present study (e.g., Cavanagh et al., 2020). Past small-scale experimental research on an adaptive system where content was available in different modalities found benefit to student learning (Mustafa & Sharif, 2011). However, that research focused on adjusting initial content presentation mode to individual learning style rather than investigating the role that availability of additional modalities may have played. Despite great interest in investigating how content in different modalities might be presented to students in e-learning systems (Khamparia & Pandey, 2020), research has not yet evaluated student learning outcomes connected to such presentation alternatives. Thus, even though adaptive systems may be designed to facilitate use of multiple modalities, the effect of doing so in them remains unknown.

Modality has been included in prior research about UDL, although not as a sole research focus. When considering content presentation overall, Orr and Hammig (2009) searched specific peer-reviewed journals from learning disability and higher education fields, between 1990 and 2008, explicitly excluding K-12. Focusing only on quantitative or qualitative empirical articles, they found 38 with research pertaining to UDL and learning disabilities. Of those, 10 contained a theme of multiple means of presentation. Thus, although studies typically do not focus solely on presentation, explicit inclusion is fairly common.

Several studies have shown support for positive student outcomes associated with universal design overall. For example, University of Minnesota faculty ran several studies of UID that found positive results for students, including higher grades and lower need for accommodations (Evans et al., 2017). Four of 80 abstracts reviewed by Mangiatori and Serenelli (2013) included assessing student academic improvement, providing general support for expectation of positive learning outcomes for UDL practices, though apparently none explicitly studied providing options for perception. Also, since these authors only reviewed

² I use representation (the official wording used by CAST) and presentation interchangeably.

abstracts and did not list the articles reviewed, these studies' quality is unknown. In a K-12 through postsecondary meta-analysis by Capp (2017) analyzing 18 pre- and post-test UDL intervention studies published between 2013 and 2016, UDL proved effective overall at improving the learning process (i.e., positive effect size reported). Two of the quantitative studies found positive effects for providing multiple means of representation via student perception self-reports. One of these compared pre/post-questionnaires of almost 400 introductory psychology students' perceptions before and after faculty teaching their course received UDL training, also comparing these results to over 200 control group students whose faculty did not receive the training (Davies et al., 2013). The other study used a 60 student convenience sample responding to a pre/post-questionnaire for one department's redesigned program study guide (Tzivinikou, 2014). Thus, studies connecting the efficacy of UDL practices to course outcomes remain scarce.

Modality Representation and Cognitive Science

The idea that humans process sensory input through multiple channels, including visual and verbal channels, has substantial evidence (Mealor et al., 2016). Likewise, the idea that human working memory has dual channels for these two pathways has years of experimental support (Mayer & Moreno, 1998). Existence of these theorized dual channels supports the idea that their use may connect to learning outcomes (Clark & Paivio, 1991; Mayer, 2001). Multimedia research tends to investigate simultaneous use of these modalities, reserving investigation of sequential modality use to the control situation (Mayer, 2001). I posit more remains to be learned about the benefits of sequential use than is currently known, however. While general multimedia research has found benefits of simultaneous presentation for learning certain content types, other research has shown neurodiverse individuals may process multimedia differently, calling into question a one-size-fits-all approach to multimedia design that assumes combining media benefits all learners in similar ways (Beacham & Alty, 2006; Wang et al., 2018). Multimedia research also tends to investigate fairly short chunks of content (e.g., a single sentence or short explanation), typically shorter than the 20-minute learning activities I study (Mayer, 2001). I draw upon cognitive science and multimedia research to support studying use of multiple modalities in sequence as well as in combination to assist struggling students with their learning.

Mayer and Massa (2003) investigated the idea that people fall into visual or verbal learner categories, and their factor analysis supported a visual/verbal distinction based on spatial ability differences, cognitive style differences, and multimedia learning preferences. Different neural information processing pathways have also been shown to operate for people with visual and verbal cognitive styles in MRI brain scans (Kraemer et al., 2009). Additionally, sensory input handling has been found to correspond to cognitive style preference by changing nonverbal information to verbal coding in the brain, for example (Kraemer et al., 2014). This supports the theory that dual pathways bring information from our senses to the point of long-term memory integration (Mayer, 2008). Additionally, memory has been found to be as good for verbal information of paragraph length presented in either an auditory or visual (textual) modality (Morris et al., 2015). It thus seems plausible that using these dual channels in sequential learning as studied here may aid long-term memory and associated learning performance requiring retention.

Additionally, “choices made within the context of an authentic learning scenario” have been found distinct from preferences expressed on questionnaires (Mayer & Massa, 2003, p. 839). This suggests a difference between innate visualizer/verbalizer cognitive style and expressed preferences when learning. Likewise, experimental brain imaging research with 20 people suggested an individual’s cognitive strategy in a given situation may be inconsistent with their questionnaire-determined cognitive style preference (Kraemer et al., 2017). This suggests that self-assessed learning style along the visual/verbal dimension may not correspond to the modality that works best for an individual when learning. This conclusion supports the idea that offering adaptive system content in different modalities may provide students benefit.

Taken together, this research suggests people use a variety of cognitive processes while learning, sometimes inconsistent with their preferred cognitive style. Combining different approaches may therefore be beneficial when individuals have a difficult time grasping information in the first way shown.

Context

The study setting is a well-established private, women’s institution in the Northeast. Older than the average four-year college-aged student coming from high school, these students typically juggle family and work responsibilities in addition to school. While such non-traditional students have frequently been underserved by higher education overall (Kazis et al., 2007), this institution supports and encourages them (Anderson & Bushey, 2017, Manly, 2023). Staff have continuously explored ways to structure course experiences and utilize data to better support student success.

The three-credit undergraduate courses studied are taught in an accelerated format in a variety of degree programs. At this institution, a semester contains three sequential subterms of six weeks each. This format allows students to take multiple courses during a subterm or multiple courses during a semester by taking one course at a time across several subterms. During the year, courses are offered across six subterms, or three times during each semester. This accelerated format facilitates working students focusing on one (or two) courses at a time while still completing multiple courses in a semester.

The technology for all courses combines a learning management system for discussion and overall course interactions (e.g., weekly assignments and grades) with an adaptive learning system for content presentation and learning mastery level formative assessments based on multiple choice questions. All courses studied had been redesigned over the prior three years using a team-based course design process utilizing Open Educational Resources (OER) within the adaptive system. This process is informed by the Quality Matters (2020) online course design rubric and annotations. Each course’s content is formatted utilizing a similar structure allowing the redesign team to code activity modality as a backend data field within the adaptive system.

Course design incorporates aspects of UDL as integrated in the Quality Matters (2020) rubric. This rubric, which aims to ensure high online course quality, guides strategies undertaken to improve student success. The rubric specifically encourages practices addressing multiple means of representation, and generally encourages following other aspects of UDL in course design (Robinson & Wizer, 2016). This aligns with arguments for broad use of

UDL as a design strategy beneficial for diverse learners (Bradshaw, 2019; Tobin & Behling, 2018). UDL adoption should go further than preserving a status quo of what constitutes good teaching that has not served some students well (Edyburn, 2010). Here this means widespread redesign of courses that incorporated UDL principles as well as concern for web accessibility in line with a mission-driven desire to improve educational success for students typically marginalized in higher education. The inclusion of multiple modalities for learning course content is a deliberate design choice made by the institution across the courses studied. Alternate paths for learning content through different modalities are part of standard course design. In an approach consistent with universal design principles around providing alternatives for perception, the adaptive learning system encourages students showing signs of struggling to pursue paths using alternate modalities until they achieve successful content mastery. To illustrate the nature of the content modalities studied, I explain an example from an introductory English course that is the second of a sequenced course pair. During an early week in this course, students are expected to gain competency in skills that would support their approach to writing. A structured activity sequence takes students through concepts needed to develop competency in the targeted skills, as in the following example of a sequence entailing three connected activities.

In the first activity, students learn about choosing a topic to write about. By default, this activity's content is presented as text. When they sufficiently master the concepts covered in the activity, as demonstrated by achieving at least 70% on a series of three to five multiple choice questions, they are allowed to progress to the next activity about how to write a thesis statement. If they show signs of struggling by not achieving at least 70%, they are redirected to review the material and a recommendation is made to view the material in a different modality. If the original content had been presented as text, they would typically have had the option to view a video if they were struggling to learn the concepts covered in the activity.³ For some activities, content in additional modalities including audio, interactive exercises, or an intentionally designed mixture of content types would also have been available in addition to text and video. Thus, students potentially had access to up to five different types of modalities for each activity, though the most common number of available modalities was two. Additional available modalities would have been accessed in a similar manner if needed by the student. Mastery of the second activity on writing a thesis statement, again demonstrated through responses to a few multiple-choice questions, then brings the student to the final activity of this sequence on writing a proposal.

In each activity, required concepts can be presented in multiple ways (i.e., different content modalities), as crafted by the course development team. As students progress through the course, their knowledge score based on the questions answered, time spent actively working on the activity, and the modality used are recorded for each activity. If they repeat an activity in a different modality, that also is recorded. This type of learning path with multiple modalities is created by developing alternative activity content for each learning objective utilizing OER to the extent possible to reduce costs for students. Depending on the subject and course, additional assignments, quizzes, and projects are also assigned and graded, as well as required weekly online discussion participation. The result is a robust dataset with ongoing measures of student action and knowledge captured throughout each

³ As will be noted later, information about the recommendations made by the adaptive learning system is considered proprietary vendor data and is not available for analysis.

week of a course. This allows analysis of student utilization of more than one modality when learning.

Data

The analysis sample includes 1278 women undergraduates enrolled in 283 sections of 51 online courses taught during the 2018–2019 academic year. These courses span 14 subjects, including sciences, social sciences, humanities, and professionally oriented courses. Student performance data allows study of the impact of using multiple modalities for representing course content on course-related success.

Several technical features facilitate the collection of these performance data, including a data warehouse and student anonymization. Student-level information is gathered from multiple campus systems and combined into a data warehouse, including data from the learning management system, adaptive platform for course content and formative assessment, and administrative student information system. Student information was anonymized prior to the researcher having access, addressing privacy concerns for this secondary data analysis.

Data were collected across multiple instances of all courses using the adaptive system during the 2018/2019 academic year. Each six-week course is broken down into learning activities each anticipated to take about 20 min, with approximately 5–15 activities per week in the adaptive system. This results in 199,396 cases for analysis.

A student's prior knowledge of the upcoming content is assessed each week and a starting knowledge state score is assigned. Knowledge of the content covered in an activity is also assessed at the completion of that activity. (While information is captured in the system about the knowledge score progression across learning attempts for each activity, to facilitate a change score analysis, the beginning and ending scores for each activity are utilized.) Information is captured about when and for how long students worked on the activity, as well as any activity repetitions and the modalities utilized each time. These features make these data well-suited for an aggregated analysis of modalities and learning across multiple courses.

Data are analyzed for each student at the activity level across all courses. Each activity instance completed by each student is given its own data row, with variables identifying whether a second modality was used at any point during that activity's completion along with the student's knowledge gain for that activity. Each activity typically has three to six content sections including a short introduction, a long section where most of the content is presented, and a summary. Questions assessing formative understanding are also asked. Sometimes at the end of a main content section, the student is asked if she would like to view alternative content. If she chooses to do so, she is often presented with content in an alternative modality, such as video if the main content is presented as text. Since the research question aims to identify an overall connection between use of more than one modality and learning gains, aggregating the data in this way across courses is sufficient.

Variables

This section explains the outcome, primary independent variable, set of independent conditioning variables, and handling of missing data. The Online Resource describes these variables' operationalization.

The outcome is change in knowledge score. An initial weekly knowledge score is assigned after determining the student's prior knowledge of that module's concepts, and exit assessments occur at the end of each 20-minute adaptive learning activity. Knowledge gain for an activity is calculated as the difference in a student's knowledge score before and after going through that activity.

The primary concept of interest is use of multiple content representations. Up to five alternate paths for learning content through different modalities are designed into each learning activity in each course. Use of multiple content representations is operationalized as student use of any second modality of content representation for a given 20-minute adaptive learning activity (from the full information about all attempts at that activity). While not all activities had the same number of modalities available, most had at least two modalities, making this treatment operationalization relevant for the greatest number of activities possible.

Two additional course-related variables are considered for model inclusion: (1) the amount of time spent on the activity since time on task may impact learning, and (2) the combination of year and term for the course since content may have been updated between terms but not during terms per institutional policy. However, the final analysis model excluded these conditioning variables, as explained below.

Missing data are not a pervasive problem. Data are only missing on the dependent variable, which could be because of either a missing starting or ending knowledge state score. When students work on the initial assessment at the beginning of the week that determined their starting knowledge level for that week's material, this activity legitimately has no beginning knowledge state score, and since this initial assessment activity is not associated with modality use while learning content, these cases are dropped from all analyses. Some students elect not to or are unable to complete the ending formative assessment after working on an activity, and this results in missingness for the ending knowledge score in 22.4% of cases. This type of missingness is expected due to the work and life demands of these non-traditional students. Given that the student has no end score in this situation, these cases are also dropped from analysis. This left analysis cases only where the student has both a beginning and ending knowledge state score.

While I acknowledge that observational research always has potential for bias due to unobserved and unknown selection effects that could be associated with missing data, sources of such bias are not anticipated here, as it seems probable that random life events led to the missing ending score. However, if struggling students are more likely to have given up and not completed the ending assessment (or to have dropped the course altogether), that might positively bias results. It is also possible though, that such missing data came from students who completed the activity with a sufficiently high score to continue along the activity sequence, but who chose to review material without completing another assessment, potentially negatively biasing results. As a sensitivity analysis utilizing all available information about these students, missing data for the ending knowledge score were han-

dled via multiple imputation (Manly & Wells, 2015), with similar analytical results (see Online Resource), lending credibility to the conclusions drawn.

Methods

Descriptive Analysis

Given the range of courses studied, I begin by breaking down the number of students, activities, and uses of multiple modalities seen across different fields of study to gauge the spread of the data across fields. Calculating means and standard errors for the analysis variables offers a descriptive sense of the data. (See Online Resource for a correlation matrix.)

Additionally, I compare demographic differences between groups that did and did not use a second modality. This allows me to investigate the potential for threats to validity caused by confounding effects of latent variables that might have caused systematic differences in outcomes of interest between groups. While there is not much that could practically be done if such problematic latent variables are unobserved, investigating systematic differences in who chose to use multiple modalities allows me to probe for potentially problematic areas that might warrant future investigation.

Aspects of the data are investigated that relate to the nature of these panel data as well. These include panel balance, amount of variation within subjects, and whether potentially problematic time-related trends were discernable.

Panel Data Analysis

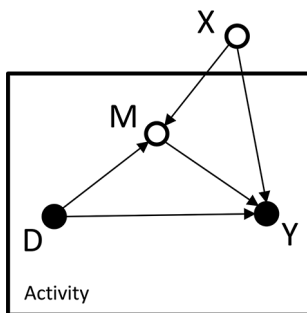
I utilize both associational and causally oriented approaches to statistical inference while addressing the clustering in the data. Investigating probable causal effects offers a particularly important and often overlooked direction for higher education research that has become increasingly possible given the more nuanced individual learning data now available through online learning systems such as those used in this study (Schneider et al., 2007). After beginning with a regression analysis to gauge the basic relationship between use of multiple modalities and knowledge gain, I explore several causally-oriented modeling approaches.

To more fully understand the relationships in the data, my approach utilizes causal graphical modeling (CGM) to represent alternative causal hypotheses that might be investigated and determine which to pursue further (Pearl, 2009). Using CGMs to represent alternative structures facilitates investigation of causal effects by aiding my modeling choices. CGMs represent random variables as nodes and causal relationships between random variables as uni-directional causal arrows between those nodes. When necessary, bi-directional arrows can also be used to indicate latent confounding. CGMs are explicit about the direction of causation whereas those relationships are either implicit or unclear in many other types of models (e.g., structural equation models). The pattern of connections among random variables that is asserted in the model directly implies marginal and conditional independencies that can be tested with data. Knowledge of the data generating process can be used to constrain the potential space of possible CGMs. For example, theory, prior research, and knowledge of time ordering can be used to infer the existence or direction of causation. In

addition to the conceptual benefits CGM can provide for developing and understanding models based on subject-matter knowledge, with a large enough dataset, relationships can be learned algorithmically from the data (Pearl, 2009; Spirtes et al., 2000). This is done here, given that the dataset had almost 200,000 observations. Known logical relationships provide constraints on this learning process to speed the processing and ensure the resulting model conformed to reality, with logical characteristic-based and time-based relationships being reflected properly. This process of model-building and testing is conducted iteratively and flexibly to determine the most appropriate model for subsequent analysis.

Using this CGM-based approach, I begin by representing the variables I expect to be related, including the use of multiple modalities (treatment, D), time spent on the activity (a potential mechanism, M, through which use of multiple modalities may have operated), the year and term to reflect possible changes in the curricular material (exogenous control, X), and the knowledge state gain score (outcome, Y). The structure of connections between these variables is also learned from the data through several Bayesian network structure learning algorithms using the *bnlearn* R package and the results are compared. These include structure learning algorithms that are constraint-based (grow-shrink, PC, and incremental association), score-based (Tabu and hill climber greedy search), and hybrid (two-phase restricted maximization, max-min hill climbing, and hybrid HPC). The models learned are constrained by prior knowledge about temporal-based relationships as well as the assumption that static characteristics (e.g., year and term) will not be predicted by other variables. After learning the edges representing relationships between these variables (i.e., nodes) from the data, the conditional probabilities of the nodes are learned by the algorithm. Testing found that the network structures from the different algorithms belong to the same equivalence class, which means these learned models from each algorithm imply the same set of conditional independencies. The resulting model of the underlying data generating process, shown in Fig. 1, indicates D and X were independent in the data.

Inspecting this model led to the conclusion that neither X nor M should be included in the analysis model. In this study, I am interested in the overall effect of use of multiple modalities on knowledge gained by the students, not specific mechanisms that might partially explain that effect, although exploring potential mechanisms, such as time on task or task repetition, could be investigated in future research. While year/term could have been included to increase precision of the effect estimate by reducing variance in the out-



Where:

D = Use of multiple modalities (treatment)

M = Time spent on activity (mechanism)

X = Year/term (exogenous)

Y = Knowledge state gain score (outcome)

● = in panel data model analyzed

Fig. 1 Graphical model learned from data

come, this was deemed unnecessary given the very large sample size. Additionally, year/term is not needed to estimate the treatment effect under a model without M. An additional exploration to learn the model when including many more of the variables that were in the full dataset did not reveal any available variables that cause both use of multiple modalities and knowledge gain (i.e., parent variables that are common causes in the language of CGMs) that ought to be included in the model. This means that whatever causes a student to use multiple modalities is either not relevant to include given my research design (such as student-level variables) or not observed and therefore not amenable to empirical investigation at this time (such as course-related variables like quality of the material or recommendations made to the student by the adaptive learning system to review the material in another modality). My conclusion from this model exploration is that the most appropriate model given my research question and the directional relationships learned from the data is a very simple panel model with only treatment and outcome, taking the clustering by student into account. Thus, I assume that the data generating process could reasonably be modeled utilizing a clustered regression analysis based on this simple graphical model (see the black dots in Fig. 1). I compare results from ordinary least squares (OLS) and panel data analyses conducted as follows.

The OLS regression is adjusted for clustering by student using Stata's *regress, vce(cluster id_student)* command (Cohen et al., 2003). To confirm the appropriateness of regression for the continuous dependent variable of gain in knowledge state score across a single activity, I verify that regression assumptions are met sufficiently. I also find no potentially problematic outliers.

I probe the causal connection between treatment and outcome using a panel data analysis with Stata's *xtreg, fe* that accounts more appropriately than OLS for the clustering of the data within individuals (Cameron & Trivedi, 2009). I take a short longitudinal approach, looking at change within student from before to after each learning activity expected to take approximately 20 min. The longitudinal nature of these data facilitates calculation of a change score across these two consecutive time points, and so a panel data analysis is appropriate to estimate the causal effect of use of multiple modalities (Hsiao, 2014). This quasi-experimental approach is known to econometricians as "a panel data variant of a difference-in-difference model" (Morgan & Winship, 2015, p. 364). Given that I have such longitudinal data from many students over courses each lasting six weeks, I estimate the effect across all activities to investigate an overall effect posited to be observable across heterogeneity in course settings and activity types.

Limitations

Several features of the data and method should be noted when interpreting the results. Although the sample likely contained many students who have disabilities, their number is unclear. This lack of clarity limits conclusions from these results for students with disabilities specifically. However, the sample's atypically low rate of official course accommodations (0.6% compared to 19% nationally; Snyder et al., 2019) may be due to the intentional design of these courses incorporating UDL principles and being guided by the Quality Matters rubric for online course design (CAST, 2018; Quality Matters, 2020). That is, students who may have felt the need to receive accommodations in other circumstances may not

have needed them for these courses. Alternatively, while it is possible that few students with disabilities chose to attend this institution in the first place, past research indicates that many students with disabilities choose not to disclose their disability in college for a number of reasons even if they had accommodations earlier in their education, and many do not know that such supports exist (Gierdowski, 2021; Newman & Madaus, 2015). Additionally, pursuing updated diagnosis and arranging for accommodations can be time consuming and expensive, and such costs may have been perceived as prohibitive, particularly for students with jobs and families who may not have much time flexibility to pursue the required process (Fox et al., 2021). Unfortunately, the reason is not possible to distinguish from available data.

Given the very large amount of data employed (almost 200,000 cases), significance tests are nearly meaningless, as even very small effects can be significant with enough data. Because of this, to aid interpretation, confidence intervals are reported to indicate estimate variation and effect sizes are emphasized.

Additionally, a potential issue with the panel data approach is that treatment assignment may have been “fuzzy” since students who received a recommendation to use a second modality might not have followed that advice. Unfortunately, it is not possible to obtain data about the recommendation offered by the adaptive learning system to students, as the adaptive learning vendor considers this proprietary information. This limitation of the present research could be addressed in future research where such data became available by the vendor.

Results

Of the 1278 students in the sample, many took courses during both Fall and Spring sessions. As shown in Table 1, 2566 learning activities were engaged in by these students. Almost 200,000 instances of activities with modality data were logged across the humanities, professional studies, math and sciences, and social sciences, and more than one modality was used over 100,000 times (58% of cases overall), again spread out by field.

Table 2 shows that across all activities, the mean knowledge improvement was 0.131 (on a 0–1 scale). Students used more than one modality while working on an activity 58% of the time. On average, students spent about 7 min (0.124 h) on an activity, and data were spread reasonably evenly between the two semesters.

When investigating the data’s panel nature, since fixed effects estimation relied on having good variation within subjects, a variance decomposition was conducted which confirmed sufficient variation existed. The panel was unbalanced, with varying case numbers

Table 1 Fields of study, students, activities, and modality use

Field of study	Students	Activities	Activities engaged by students	Times > 1 modality used	% > 1 modality used
Humanities	745	594	56,821	31,301	55%
Professional studies	557	1168	59,466	41,074	69%
Math and sciences	417	343	40,682	12,061	30%
Social sciences	585	467	42,427	31,187	74%
Total	1278	2566	199,396	115,623	58%

Table 2 Estimated means and standard errors of the estimates

Variable	Mean	SE
Knowledge state gain across activity	0.131	<0.000
Use of > 1 modality?	0.580	0.001
Hours spent on activity	0.124	0.001
Fall 2018	0.530	0.001
Spring 2019	0.470	0.001
Observations	199,396	

Table 3 Average marginal effects for knowledge gain across an activity – clustered ordinary least squares regression and fixed effects panel data analyses

Variables	(1) Regression with student clustering (SE)	[CI]	(2) Fixed effects panel data analysis (SE)	[CI]
Use of > 1 modality?	0.069*** (0.003)	[0.061, 0.076]	0.049*** (0.003)	[0.043, 0.056]
Constant	0.091*** (0.003)	[0.084, 0.098]	0.103*** (0.001)	[0.099, 0.106]
Observations	199,396		199,396	
R ²	0.023		0.009	

Note Standard errors in parentheses. Confidence interval in brackets. Low R² values do not pose a problem since this treatment effect study is not focused on explaining outcome variance

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

for students, because different courses had different numbers of activities and students could choose whether to complete them. 89.4% of unique students had activities where they only used one modality, while 98.5% of unique students had activities where they used more than one modality. Thus, many students engaged in both kinds of approaches for different activities. Students engaged in more than one modality across 80.6% of unique activities. Within courses, data were clustered for each student. Checking time-series plots showed no potentially problematic discernable trends over time. Testing for heteroskedasticity suggested that using clustering by student was indeed appropriate for these data ($\chi^2 = 868$, $p < 0.001$). A simple cross-validation check splitting the data by semester confirmed that results in each semester were similar to those presented below. Finally, a robust Hausman test confirmed the appropriate use of fixed effects for these data ($\chi^2 = 39$, $p < 0.001$). Thus, my analysis focused on a fixed effects panel data model.

As shown in Table 3, the average marginal effect of use of more than one modality to learn the content in an activity was 0.049 when calculated with a fixed effects panel approach (model 2) accounting for student-level factors that might influence results. Clustered regression results (model 1) are presented for comparison. The panel coefficient (0.049) is equivalent to a standardized effect size of Hedges' $g = 0.224$ (see Table 4). This can be interpreted as a reasonable effect size for education since Cohen's labeling of 0.20 as small and 0.50 as medium "can be misleading in educational policy contexts, in which effect sizes of 0.20 or smaller are often of policy interest" (Hedges & Hedberg, 2007). Recent guidance for educational interventions considers effects over 0.20 to be large (Kraft, 2020), though some in higher education would argue for slightly larger values (Mayhew et al., 2016). The effect found corresponds to an improvement index of +8.9 (above 50th

Table 4 Hedges' *g* effect sizes corresponding to analysis models in Table 3

Model	Comparison		Treatment		Impact			Effect size
	Mean	SD	Mean	SD	Estimate	SE	p	
(1)	0.091	0.206	0.160	0.228	0.069	0.003	<0.000	0.314
(2)	0.091	0.206	0.140	0.228	0.049	0.003	<0.000	0.224
<i>N</i>	83,773		115,623					

percentile), which is equivalent to a comparison student improving from the 50th to the 59th percentile (What Works Clearinghouse, 2020).

The sensitivity of results to choices made during analysis were probed through several additional analyses (see Online Resource). Substantively similar conclusions to those presented were drawn when using two alternate operationalizations of the “mixed” treatment category, running OLS regression including additional covariates, multiply imputing missing values on the dependent variable, and adjusting for clustering by activity. Additionally, learning the structure of more complex models through a Bayesian network approach did not suggest potential confounders that should be included for an analysis of the overall effect, although it did suggest possible mediating factors that could be investigated in future research.

To check for balance in demographic and prior educational factors across treatment groups, I tested for differences between cases that did and did not use multiple modalities in race/ethnicity, Pell grant status, age, withdrawals and failures in the prior semester, prior GPA, and the number of credits transferred in when the student entered the institution (see Online Resource). Finding only a significant difference by age, I probed further and found a difference between students under and over the median age of 31, although there was no difference within each of these groups. This suggests future research might explore differences in modality use between younger and older students.

Discussion and Implications

My analysis found a medium-large, educationally important effect of using multiple modalities on the knowledge gain students exhibit across a learning activity. This work extends calls to scientifically validate aspects of UDL, supporting guidance to provide flexible options for perceiving content as a way to deeply connect students with the material they are learning (Edyburn, 2010; Rao et al., 2014). On average, use of more than one modality predicted a 0.05 knowledge score increase on a 0–1 scale across a learning activity over students using only one modality. This corresponded to a student improving almost 10 percentiles above the activity median, a meaningful boost. This is in line with expectations that providing content in multiple modalities will assist student learning (Rose, 2001). These results make one of UDL's benefits for formative student learning outcomes concrete, offering a contribution to the universal design literature, which has been lacking in efficacy studies (Cumming & Rose, 2021; Roberts et al., 2011). Overall, the results of this study support UDL's claim that providing multiple means of representing content will be beneficial, quantifying that benefit for women in the adaptive learning context studied.

Given the large sample, in determining confidence that the result indicates a real effect, I also investigated the amount of bias it would have taken to switch from a significant to non-significant finding (Frank et al., 2013). I note that the effect would have needed to be

biased by 89.95% to invalidate the inference. Alternately, it would have taken a confounding variable correlated at 0.199 with both treatment and outcome to invalidate the result. Such a correlation with the outcome would have been stronger than the outcome's correlation with either the treatment (0.153) or the amount of time spent on the activity (0.039). This gives confidence that the result is quite strong, even considering the large sample size. The volume of data is a study strength, while also being large enough to warrant emphasizing effect size interpretation over statistical significance.

While caution is always warranted when making causal claims, the panel nature of the data means person-centered variables that are difficult to measure and often confound observational studies should not bias these results. That is, in typical regression modeling accounting for clustered data, collecting data about personal background factors may be challenging or practically impossible. Such factors could include motivation, personality-based predispositions, or prior experiences that serve to increase engagement with the material. While observable characteristics can be measured and models adjusted appropriately, potential exists for unobserved characteristics to introduce bias. A panel approach essentially allows a given student to act as her own comparison, automatically adjusting for person-related factors so they will not confound conclusions drawn. Non-student factors may still have biased the results, such as the quality of either material, course design, implementation of UDL principles, or instructor teaching. However, use of the Quality Matters (2020) rubric by the institution in the development of these courses supports the assumption that such quality measures were held constant in this analysis, supporting a causal interpretation of the results. It is also possible that clustering effects at the course level may be relevant, but since some students took multiple courses, the necessary cross-classified analysis needed to investigate this is beyond the scope of this study and is left for future research to investigate. Overall, the panel approach held notable strength for a person-oriented outcome as studied here, particularly when coupled with approaches to ensure baseline course quality, even while future research about possible alternative explanations beyond student-level factors remains warranted.

Additionally, both a data science-oriented approach of learning the model from data and multiple sensitivity analyses probing the influence of a variety of choices made during analysis suggest confidence in the conclusions drawn. Although online course-taking during remote learning at the height of the COVID-19 pandemic may not have reflected the voluntary nature of the choice to study online by the students in the earlier time period studied here (Hodges et al., 2020), it is reasonable to assume future students will again choose online courses on a voluntary rather than forced basis. Accordingly, these conclusions have expected relevance going forward. Thus, I claim with reasonable conviction that non-traditional, undergraduate women students of differing ability levels taking online courses benefit from the opportunity to learn content by utilizing multiple modalities across a range of humanities, professional, social science, and scientific disciplines.

Alternative Explanations

Multiple causal paths may underlie the improvement seen in learning gain given use of multiple modalities and such alternative explanations are worthwhile to consider when interpreting the results of this treatment effect study. For example, learning might improve when students repeat the activity, giving them more exposure to the material. Although the impor-

tance of time spent learning might have face validity and has generally been considered good practice to encourage (Chickering & Gamson, 1987), prior research has sometimes found positive (Wellman & Marcinkiewicz, 2004) and sometimes negative (Greenwald & Gillmore, 1997) relationships between time-related factors and student achievement, so the potential influence on the present study is unclear. Preliminary investigation of possible alternative explanations for the results reported here such as time on task and activity repetition did not appear to explain away the effect of use of multiple modalities when included in preliminary sensitivity analyses, reinforcing confidence in the claim of a positive effect that may be durable even when considering potential mediators. However, future mediation-focused research could investigate the extent to which these and other factors may be operating in concert to aid students' learning. From a causal perspective, such potential alternative explanations should be researched to determine the extent to which they are also important in understanding UDL and modality use.

Alternatively, an argument might be made that a particular modality is simply "better" at conveying certain content. For example, a faculty member who learned a concept in a particular way may believe that way to be "the best." However, UDL principles "[reflect] the fact that there is no one way of presenting information or transferring knowledge that is optimal for all students" (Rose et al., 2006, p. 137). Based on the reality that perceptual capabilities differ between individuals (Mealor et al., 2016), I would not expect that certain material would be found to be most effectively conveyed through a particular modality for all students. If future discipline-specific or course-specific research found evidence to the contrary, this would point to an alternative explanation that might confound the results of the present study and challenge this foundational UDL principle, though this seems unlikely.

As another candidate cause to consider, it is possible that the second modality used by students was better suited to their learning needs. Students may not have been guided to an optimal initial choice for conveying content by the adaptive system's default learning path. Recognizing this possibility, over time, the adaptive system notes which content presentation mode works better for a given student based on their performance and will begin presenting material in that modality first when alternative content is available (Cavanagh et al., 2020). Although the system was implemented for too short a time to expect confounding from adaptation of the initial modality presented and few students would have known how to change the default initial modality, future research could investigate possible order effects. Based on results refuting the matching hypothesis in the learning styles literature (Cuevas, 2015; Pashler et al., 2009), I would not expect that matching students who prefer a given modality type with material presented solely in that modality would improve learning. It is less clear whether using particular combinations or sequences of modalities might be beneficial given that such combinations may tap into the different brain regions people use when processing visual and verbal information (Kraemer et al., 2009). It is also unclear whether any such combinatory effect might differ for students who report particular learning preferences, such as visual or verbal (Mayer & Massa, 2003). The present results suggest that future research investigating specific sequences of modality use would be warranted.

As another possible cause, opportunities to make choices have been considered a component of student agency leading to improved academic performance (Jääskelä et al., 2021). Here, the agency that comes with freedom of choice to pursue different modalities may have been operating to aid students' learning. While this could be investigated in future research, to put this and other possible alternative causes in context as already noted, such a poten-

tially confounding variable would need to have had a 0.2 correlation with both treatment and outcome to nullify the treatment effect. In that case, such a correlation with freedom of choice would have been stronger than either's correlation with time on task.

Future Research

The results suggest numerous additional intriguing directions for future research. The possibility exists that factors such as motivation to earn a high grade may moderate the results. That is, the institution's learning design team is aware that some students who are very motivated to earn a high grade will repeat activities over and over until they earn high grades on every activity. This anecdotal information is in line with prior research on agency, self-efficacy, and high performing students that has found motivated students do better academically and competitive students will work hard to achieve a high grade (Alkış & Temizel, 2018; Ayllón et al., 2019; Baumann & Harvey, 2021). In the present research, a latent factor for such grade-based motivation was used in a regression sensitivity analysis (with no substantive difference in result; see Online Resource), but was not employed in the panel analysis since that factor was constant for a given student. However, future research could consider stratifying the sample by a measure of grade motivation to investigate the possibility that use of multiple modalities may operate differently for students with higher or lower motivation to achieve a high grade. In a similar vein, other potential moderating factors such as prior academic achievement could be investigated to gain a fuller picture of the circumstances under which use of multiple modalities makes the biggest positive difference for students.

Research taking a more nuanced look at the specific modalities used by students could also investigate whether use of more than two modalities offers benefit (e.g., in a dosage analysis). Existing theory about dual-channel visual/auditory processing suggests that the largest cognitive difference may come from use of modalities offering complementary visual and auditory sensory input (Mayer, 2001). From this standpoint, a third, fourth, or fifth mode that uses different combinations of sensing and cognition to process (e.g., video involves both visual and auditory elements) may duplicate the sensory input of either a single visual- or auditory-based presentation. It remains unclear whether the impact of use of a second modality is related primarily to a dual input distinction (i.e., eye and ear), to a dual processing distinction (i.e., visual and auditory), or to dual-channel pathways within working memory overall (Kraemer et al., 2009, 2014; Mayer, 2008). Given these dualities, the benefit of using multiple modalities may primarily be a benefit of using at least a second modality. Consistent with this supposition, a preliminary look at dosage suggested the biggest benefit may appear after use of any second modality. However, not all courses studied had content in more than two modalities available, so this characteristic of the data may have had a confounding influence on these preliminary dosage explorations. Future research should distinguish the limits and causes of dosage effects further.

Several other directions left unexplored by the present study could also be targeted in future research. For example, the effect may be stronger for some courses or subjects than others. The timing of content presentation in alternate modalities by the adaptive system might matter, involving analysis of recommendations made to struggling students to use another modality. Students dropping the course may have had more difficulty with the material, so an intermediate outcome of course withdrawal could be investigated to gauge pos-

sible effect attenuation. The number and type of modalities offered for a given activity may matter as well, presenting a potential confounding influence which could be researched, informing future course design.

Past research investigating learning through brain mechanisms involving multiple sensory pathways to memory supports the idea that the present results may have wider applicability (Mayer, 2008; Mayer & Moreno, 1998). When investigating the effect of the simultaneous presentation of media utilizing dual-channel sensory pathways in complementary fashion, multimedia research has found benefit in utilizing both visual and auditory sensory modalities when learners are to remember and integrate content information (Mayer, 2001). Although such work has focused on simultaneous presentation of multiple media, the content presentation in the present study also makes use of more than one sensory channel for learning, but primarily for consecutive presentation. Although multimedia may be more effective than single-mode presentation for some types of students learning some types of content, students with certain types of disabilities, such as dyslexia, may encounter difficulty comprehending material presented simultaneously in different modes due to the required cognitive load (Beacham & Alty, 2006). So, while the sequential type of presentation studied here is perhaps not as efficient a method of comprehending material as a full multimedia presentation for some individuals, it deliberately offers choices and alternative learning paths to students, giving them agency to utilize what works for them. Additionally, the adaptive learning system does not overwhelm students with too many options initially, keeping the cognitive load down, which can otherwise challenge some neurodiverse students depending on how options for multiple modalities are implemented (Kohler & Balduzzi, 2021). The results of the present study reveal potential benefits of combining ideas about dual-channel processing for memory with dual coding for cognitive load, supporting the idea that memory function is not necessarily dependent on the type of sensory input (Morris et al., 2015). That is, memory benefits that exist when utilizing both visual and auditory channels to reinforce learning appear to operate under the conditions studied here in ways that support retention of material without potentially overtaxing cognitive load, thus effectively addressing a wide range of cognitive abilities.

It seems reasonable to suspect these results may hold more broadly even while acknowledging the limits of external validity for a single research study of one institution and the need to extend this work to a variety of student populations. Given the paucity of research literature addressing the effectiveness of practices based on UDL guidelines for improving student learning (Mangiatordi & Serenelli, 2013), it would be helpful to undertake studies exploring the extent to which these results can be replicated in other settings. Additional research could be undertaken to investigate the effect on outcomes at different time scales and in different institutional contexts beyond a small women's college and beyond the online setting. The result could be confirmed for men as well as women, and for traditional age students as well as the predominantly non-traditional students studied here. While more work remains to confirm these results with students who have known disabilities, current results support a claim of broad applicability for providing content in multiple modalities. The demonstrated benefit realized by many students in this institutional context suggests that course design steps such as those taken by this institution may reduce the need for specific accommodations. Future research should distinguish between such a design-based reason and other possible explanations, such as students not seeking accommodations that would

benefit them due to perceived or feared stigma or lack of knowledge about the accommodation process.

The current research also leaves open the question of why students choose other modalities. Having identified that use of multiple modalities provides benefit for non-traditional undergraduate women, future qualitative research could interview students to investigate why they chose to use multiple modalities, what they hoped to gain by doing so, and how they perceived the benefit obtained. Better understanding student motivation to engage in working through content in different modalities may help future educators design courses that encourage the positive aspects of this practice more explicitly.

Implications for Practice

The strong evidence presented here for an educationally meaningful positive effect of use of multiple modalities has important implications for practice. These results provide a compelling argument that faculty development and curricular design efforts should include the UDL principle of providing multiple means of representation for course content. That is, there are demonstrable benefits for formative learning gains when students are given the opportunity to encounter course content in more than one modality. Faculty development increasingly includes exposing faculty to universal design principles, and widely used guidelines for good online development incorporate UDL ideas (Higbee & Goff, 2008; Robinson & Wizer, 2016). However, even though faculty are often aware of the need to learn about and implement UDL ideas, this does not always translate to actual implementation (Cook et al., 2009; Izzo et al., 2008). Encouragingly though, faculty who have received UDL training are more likely to include multiple means of presentation in their teaching (Lombardi et al., 2011). In line with what has been termed the “plus one” strategy for approaching UDL implementation (Tobin & Behling, 2018), identifying key material where students typically struggle and adding an alternative for learning content in a different modality for that material may be a good place for faculty to start as they add to their UDL-informed practice and work toward fully incorporating UDL concepts. This study provides clear and compelling support for making options for content available, and action to achieve this can be encouraged in faculty training.

This study also provides concrete evidence that curriculum development efforts should include making content available in multiple modalities, particularly in adaptive learning systems, because students can see an improvement index of almost +10 above the median in their learning. At the institution studied, a systematic and comprehensive approach to including multiple modalities was strategically undertaken, with a design team adding such material to over 50 courses. Such modality options can include alternate text, video, audio, interactive, or mixed modality representations of content. The benefits seen suggest other institutions would be well advised to consider devoting resources to systematically developing options for students to go through material in science, social science, humanities, and professionally oriented fields. Offering students options for how content is presented is a commonsense UDL tenet with demonstrable benefit that is straightforward for faculty and institutions to implement if they have allocated sufficient resources for implementation. Such clear opportunities to improve practice are all too rare in postsecondary education and should be a call to action.

Conclusion

This study investigated the relationship between use of multiple content representations and formative student outcomes for 20-minute adaptive learning activities in an adaptive learning system. The goal was to better understand and help confirm UDL's proposition that providing multiple means of content representation benefits student learning. This work extends knowledge about UDL in practice by identifying the effect of the use of multiple modalities on formative learning done by women undergraduates as they engaged with content for online courses across multiple fields. By combining data from several campus systems, a comprehensive within-course dataset enabled estimates of effects to be revealed through a within-subjects analysis approach. Results support UDL's claimed benefit of providing options for perception by demonstrating quantifiable learning gains for students. This suggests that time spent by faculty and course developers modifying course material to incorporate different modalities offers clear benefit to students. These results should bolster administrative efforts to direct resources, such as faculty development funding and support, toward efforts to provide content to students in multiple modalities.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11162-024-09784-9>.

Acknowledgements A previous version of this paper was presented at the Northeastern Educational Research Association Annual Conference in Trumbull, Connecticut, in October 2018, where it received the GSIC Best Paper by a Graduate Student Award. The author would like to thank Ryan Wells, David Jensen, and Ezekiel Kimball for helpful comments on previous versions of this paper.

Funding No funding was received for conducting this study.

Data Availability The data that support the findings of this study are not publicly available. The agreement with the institution to analyze the data does not allow the researcher to publicly share the data.

Code Availability Stata code used for analysis is presented in the Online Resource.

Declarations

Ethical Approval Institutional Review Board approval was secured prior to conducting this study.

Competing Interests The author has no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abell, M. M., Jung, E., & Taylor, M. (2011). Students' perceptions of classroom instructional environments in the context of Universal Design for Learning. *Learning Environments Research*, 14(2), 171–185. <https://doi.org/10.1007/s10984-011-9090-2>.
- Alkış, N., & Temizel, T. T. (2018). The impact of motivation and personality on academic performance in online and blended learning environments. *Journal of Educational Technology & Society*, 21(3), 35–47.
- Anderson, J., & Bushey, H. (2017). Journey to adaptive learning at The American Women's College. *EDUCAUSE Review*. <https://er.educause.edu/articles/2017/6/journey-to-adaptive-learning-at-the-american-womens-college>.
- Association on Higher Education and Disability (2021). *AHEAD Statement on Language*. <https://www.ahead.org/professional-resources/accommodations/statement-on-language>.
- Ayllón, S., Alsina, Á., & Colomer, J. (2019). Teachers' involvement and students' self-efficacy: Keys to achievement in higher education. *Plos One*, 14(5), 1–11. <https://doi.org/10.1371/journal.pone.0216865>.
- Baumann, C., & Harvey, M. (2021). What is unique about high performing students? Exploring personality, motivation and competitiveness. *Assessment & Evaluation in Higher Education*, 0(0), 1–13. <https://doi.org/10.1080/02602938.2020.1870930>.
- Beacham, N. A., & Alty, J. L. (2006). An investigation into the effects that digital media can have on the learning outcomes of individuals who have dyslexia. *Computers & Education*, 47(1), 74–93. <https://doi.org/10.1016/j.compedu.2004.10.006>.
- Bradshaw, D. G. (2019). Proactive design and inclusive practices: Universal Design for Learning in higher education [Ed.D., California State University, Los Angeles]. In *ProQuest Dissertations and Theses*. <http://search.proquest.com/docview/2457368678/abstract/1E24ECC00E6C4BC6PQ/1>.
- Burgstahler, S. E. (2015). *Universal design in higher education: From principles to practice* (2nd ed.). Harvard Education.
- Burgstahler, S. E., & Cory, R. (Eds.). (2008). *Universal design in higher education: From principles to practice*. Harvard Education.
- Cameron, A. C., & Trivedi, P. K. (2009). *Microeconometrics using Stata* (Revised). Stata Press.
- Capp, M. J. (2017). The effectiveness of Universal Design for Learning: A meta-analysis of literature between 2013 and 2016. *International Journal of Inclusive Education*, 21(8), 791–807. <https://doi.org/10.1080/13603116.2017.1325074>.
- CAST (2014). *UDL guidelines—Version 2.0: Principle I. Provide multiple means of representation*. Universal Design for Learning Guidelines Version 2.0. http://www.udlcenter.org/aboutudl/udlguidelines/principle1#principle1_g1.
- CAST (2018). *Universal Design for Learning guidelines version 2.2*. <https://udlguidelines.cast.org/>.
- Cavanagh, T., Chen, B., Lahcen, R. A. M., & Paradiso, J. R. (2020). Constructing a design framework and pedagogical approach for adaptive learning in higher education: A practitioner's perspective. *International Review of Research in Open and Distributed Learning*, 21(1), 173–197. <https://doi.org/10.19173/irrodl.v21i1.4557>.
- Chickering, A. W., & Gamson, Z. F. (1987). Seven principles for good practice in undergraduate education. *American Association of Higher Education Bulletin*, 39(7), 3–7.
- Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, 3(3), 149–210.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Routledge.
- Cook, L., Rumrill, P. D., & Tankersley, M. (2009). Priorities and understanding of faculty members regarding college students with disabilities. *International Journal of Teaching and Learning in Higher Education*, 21(1), 84–96.
- Crevecoeur, Y., Sorenson, S., Mayorga, V., & Gonzalez, A. (2014). Universal Design for Learning in K-12 educational settings: A review of group comparison and single-subject intervention studies. *The Journal of Special Education Apprenticeship*, 3(2). <http://scholarworks.lib.csusb.edu/josea/vol3/iss2/1>.
- Cuevas, J. (2015). Is learning styles-based instruction effective? A comprehensive analysis of recent research on learning styles. *Theory and Research in Education*, 13(3), 308–333. <https://doi.org/10.1177/1477878515606621>.
- Cumming, T. M., & Rose, M. C. (2021). Exploring Universal Design for Learning as an accessibility tool in higher education: A review of the current literature. *Australian Educational Researcher*, 1–19. <https://doi.org/10.1007/s13384-021-00471-7>.
- Davies, P. L., Schelly, C. L., & Spooner, C. L. (2013). Measuring the effectiveness of Universal Design for learning intervention in postsecondary education. *Journal of Postsecondary Education and Disability*, 26(3), 195–220.

- Dorn, E., Hancock, B., Sarakatsannis, J., & Viruleg, E. (2020). *COVID-19 and student learning in the United States: The hurt could last a lifetime*. <https://www.mckinsey.com/industries/public-sector/our-insights/covid-19-and-student-learning-in-the-united-states-the-hurt-could-last-a-lifetime>.
- Edyburn, D. L. (2010). Would you recognize universal design for learning if you saw it? Ten propositions for new directions for the second decade of UDL. *Learning Disability Quarterly*, 33(1), 33–41. <https://doi.org/10.1177/073194871003300103>.
- Evans, N. J., Broido, E. M., Brown, K. R., & Wilke, A. K. (2017). *Disability in higher education: A social justice approach*. Jossey-Bass.
- Fovet, F., & Mole, H. (2013). UDL-from disabilities office to mainstream class: How the tools of a minority address the aspirations of the student body at large. *Collected Essays on Learning and Teaching*, 6, 121–126.
- Fox, A., Hedayet, M., Mansour, K. E., Kommers, S., & Wells, R. (2021). College students with disabilities experiences with financial, social, and emotional costs on campus in the United States. *International Journal of Disability Development & Education*, 1–15. <https://doi.org/10.1080/1034912x.2021.1966758>.
- Frank, K. A., Maroulis, S. J., Duong, M. Q., & Kelcey, B. M. (2013). What would it take to change an inference? Using Rubin's causal model to interpret the robustness of causal inferences. *Educational Evaluation and Policy Analysis*, 35(4), 437–460.
- Gierdowski, D. C. (2021). Accessibility hazards on the road to college. *EDUCAUSE Review*. <https://er.educause.edu/blogs/2020/11/accessibility-hazards-on-the-road-to-college>.
- Greenwald, A. G., & Gillmore, G. M. (1997). No pain, no gain? The importance of measuring course workload in student ratings of instruction. *Journal of Educational Psychology*, 89(4), 743–751. <https://doi.org/10.1037/0022-0663.89.4.743>.
- Greer, D. L., Crutchfield, S. A., & Woods, K. L. (2013). Cognitive theory of multimedia learning, instructional design principles, and students with learning disabilities in computer-based and online learning environments. *Journal of Education*, 193(2), 41–50. <https://doi.org/10.1177/002205741319300205>.
- Hall, T. E., Cohen, N., Vue, G., & Ganley, P. (2015). Addressing learning disabilities with UDL and technology: Strategic Reader. *Learning Disability Quarterly*, 38(2), 72–83. <https://doi.org/10.1177/0731948714544375>.
- Hamraie, A. (2017). *Building access: Universal design and the politics of disability*. University of Minnesota.
- Hedges, L. V., & Hedberg, E. C. (2007). Intraclass correlation values for planning group-randomized trials in education. *Educational Evaluation and Policy Analysis*, 29(1), 60–87.
- Herold, D. S., & Chen, T. (2021). Switching from face-to-face to online instruction midsemester: Implications for student learning. *Journal of Teaching and Learning with Technology*, 10, 321–336.
- Higbee, J. L., & Goff, E. (2008). *Pedagogy and student services for institutional transformation: Implementing universal design in higher education*. Center for Research on Developmental Education and Urban Literacy. <https://eric.ed.gov/?id=ED503835>.
- Hodges, C., Moore, S., Lockee, B., Trust, T., & Bond, A. (2020). The difference between emergency remote teaching and online learning. *EDUCAUSE Review*, 27, 1–12.
- Hsiao, C. (2014). *Analysis of panel data* (3rd ed.). Cambridge University Press.
- Izzo, M. V., Murray, A., & Novak, J. (2008). The faculty perspective on Universal Design for Learning. *Journal of Postsecondary Education and Disability*, 21(2), 60–72.
- Jääskelä, P., Heilala, V., Kärkkäinen, T., & Häkkinen, P. (2021). Student agency analytics: Learning analytics as a tool for analysing student agency in higher education. *Behaviour & Information Technology*, 40(8), 790–808. <https://doi.org/10.1080/0144929X.2020.1725130>.
- Kazis, R., Callahan, A., Davidson, C., McLeod, A., Bosworth, B., Choitz, V., & Hoops, J. (2007). *Adult learners in higher education: Barriers to success and strategies to improve results* (Occasional Paper No. 2007–03). U.S. Department of Labor. <https://www.jff.org/resources/adult-learners-higher-education-barriers-success-and-strategies-improve-results/>.
- Khamparia, A., & Pandey, B. (2020). Association of learning styles with different e-learning problems: A systematic review and classification. *Education and Information Technologies*, 25(2), 1303–1331. <https://doi.org/10.1007/s10639-019-10028-y>.
- Kimball, E. W., Wells, R. S., Ostiguy, B. J., Manly, C. A., & Lauterbach, A. (2016). Students with disabilities in higher education: A review of the literature and an agenda for future research. *Higher Education: Handbook of Theory and Research*, 31, 91–156.
- Kohler, M., & Balduzzi, T. (2021). *The forgotten ones: Exploring the LXD model to address the barriers UDL creates for neurodivergent learners*. OLC Accelerate 2021 Conference, Washington, DC. <https://onlinelearningconsortium.org/olc-accelerate-2021-session-page/?session=10316&kwds=>
- Kortering, L. J., McClannnon, T. W., & Braziel, P. M. (2008). Universal Design for Learning: A look at what algebra and biology students with and without high incidence conditions are saying. *Remedial and Special Education*, 29(6), 352–363. <https://doi.org/10.1177/0741932507314020>.

- Kraemer, D. J. M., Rosenberg, L. M., & Thompson-Schill, S. L. (2009). The neural correlates of visual and verbal cognitive styles. *Journal of Neuroscience*, 29(12), 3792–3798. <https://doi.org/10.1523/JNEUROSCI.4635-08.2009>.
- Kraemer, D. J. M., Hamilton, R., Messing, S. B., DeSantis, J. H., & Thompson-Schill, S. L. (2014). Cognitive style, cortical stimulation, and the conversion hypothesis. *Frontiers in Human Neuroscience*, 8, Article 15. <https://doi.org/10.3389/fnhum.2014.00015>.
- Kraemer, D. J. M., Schinazi, V. R., Cawkwell, P. B., Tekriwal, A., Epstein, R. A., & Thompson-Schill, S. L. (2017). Verbalizing, visualizing, and navigating: The effect of strategies on encoding a large-scale virtual environment. *Journal of Experimental Psychology: Learning Memory and Cognition*, 43(4), 611–621. <https://doi.org/10.1037/xlm0000314>.
- Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, 49(4), 241–253. <https://doi.org/10.3102/0013189X20912798>.
- Lombardi, A. R., Murray, C., & Gerdes, H. (2011). College faculty and inclusive instruction: Self-reported attitudes and actions pertaining to Universal Design. *Journal of Diversity in Higher Education*, 4(4), 250–261. <https://doi.org/10.1037/a0024961>.
- Manly, C. A. (2023). When adaptivity and Universal Design for Learning are not enough: Bayesian network recommendations for tutoring. In P. D. Moskal, C. D. Dziuban, & A. G. Picciano (Eds.), *Data analytics and adaptive learning*. Routledge/Taylor. <https://doi.org/10.4324/9781003244271>.
- Mangiardi, A., & Serenelli, F. (2013). Universal Design for Learning: A meta-analytic review of 80 abstracts from peer reviewed journals. *Research on Education and Media*, 5(1), 109–118.
- Manly, C. A., & Wells, R. S. (2015). Reporting the use of multiple imputation for missing data in higher education research. *Research in Higher Education*, 56(4), 397–409. <https://doi.org/10.1007/s11162-014-9344-9>.
- Manly, C. A., Wladis, C., Hachey, A. C., & Conway, K. M. (2021). *Did COVID-19 exacerbate inequities? Course outcomes, on average and when considering students' prior online experience*. OLC Accelerate Conference, Washington, DC.
- Mayer, R. E. (2001). *Multimedia learning*. Cambridge University Press.
- Mayer, R. E. (2008). Applying the science of learning: Evidence-based principles for the design of multimedia instruction. *American Psychologist*, 63(8), 760–769. <https://doi.org/10.1037/0003-066X.63.8.760>.
- Mayer, R. E., & Massa, L. J. (2003). Three facets of visual and verbal learners: Cognitive ability, cognitive style, and learning preference. *Journal of Educational Psychology*, 95(4), 833–846. <https://doi.org/10.1037/0022-0663.95.4.833>.
- Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology*, 90(2), 312–320. <https://doi.org/10.1037/0022-0663.90.2.312>.
- Mayhew, M. J., Rockenbach, A. N., Bowman, N. A., Seifert, T. A. D., Wolniak, G. C., Pascarella, E. T., & Terenzini, P. T. (2016). *How college affects students: 21st century evidence that higher education works*. John Wiley & Sons, Inc.
- McGuire, J. M. (2014). Universally accessible instruction: Oxymoron or opportunity? *Journal of Postsecondary Education and Disability*, 27(4), 387–398.
- Mealor, A. D., Simner, J., Rothen, N., Carmichael, D. A., & Ward, J. (2016). Different dimensions of cognitive style in typical and atypical cognition: New evidence and a new measurement tool. *PLOS ONE*, 11(5), e0155483. <https://doi.org/10.1371/journal.pone.0155483>.
- Meyer, A., Rose, D. H., & Gordon, D. (2014). *Universal Design for Learning: Theory and practice*. CAST Professional Publishing.
- Morgan, S. L., & Winship, C. (2015). *Counterfactuals and causal inference: Methods and principles for social research* (2nd ed.). Cambridge University Press.
- Morris, J., Woodworth, C., Swier-Vosnos, A., Rossini, E., & Jackson, I. (2015). Relationship of sensory modality to retention of episodic memory. *Applied Neuropsychology: Adult*, 22(4), 246–251. <https://doi.org/10.1080/23279095.2014.914517>.
- Mustafa, Y. E. A., & Sharif, S. M. (2011). An approach to adaptive E-learning Hypermedia System based on learning styles (AEHS-LS): Implementation and evaluation. *International Journal of Library and Information Science*, 3(1), 15–28. <https://doi.org/10.5897/IJLIS.9000009>.
- Newman, L. A., & Madaus, J. W. (2015). Reported accommodations and supports provided to secondary and postsecondary students with disabilities: National perspective. *Career Development and Transition for Exceptional Individuals*, 38(3), 173–181. <https://doi.org/10.1177/2165143413518235>.
- Office for Civil Rights (2021). *Education in a pandemic: The disparate impacts of COVID-19 on America's students*. <https://www2.ed.gov/about/offices/list/ocr/docs/20210608-impacts-of-covid19.pdf>.
- Orr, A. C., & Hammig, S. B. (2009). Inclusive postsecondary strategies for teaching students with learning disabilities: A review of the literature. *Learning Disability Quarterly*, 32(3), 181–196.

- Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2009). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest*, 9(3), 105–119. <https://doi.org/10.1111/j.1539-6053.2009.01038.x>.
- Pearl, J. (2009). *Causality: Models, reasoning, and inference* (2nd ed.). Cambridge University Press.
- Quality Matters (2020). *Specific review standards from the QM higher education rubric, 6th edition*. Quality Matters. <https://www.qualitymatters.org/qa-resources/rubric-standards/higher-ed-rubric>.
- 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. <https://doi.org/10.1177/0741932513518980>.
- Roberts, K. D., Park, H. J., Brown, S., & Cook, B. (2011). Universal Design for instruction in postsecondary education: A systematic review of empirically based articles. *Journal of Postsecondary Education and Disability*, 24(1), 5–15.
- Robinson, D. E., & Wizer, D. R. (2016). Universal Design for Learning and the quality matters guidelines for the design and implementation of online learning events. *International Journal of Technology in Teaching and Learning*, 12(1), 17–32.
- Rose, D. H. (2001). Universal Design for Learning: Deriving guiding principles from networks that learn. *Journal of Special Education Technology*, 16(2), 66–67.
- Rose, D. H., Harbour, W., Johnston, C. S., Daley, S., & Abarbanell, L. (2006). *Universal Design for learning in postsecondary education: Reflections on principles and their application*. National Center on Universal Design for Learning. <http://www.udlcenter.org/sites/udlcenter.org/files/UDLinPostsecondary.pdf>.
- Schneider, B., Carnoy, M., Kilpatrick, J., Schmidt, W. H., & Shavelson, R. J. (2007). *Estimating causal effects using experimental and observational designs*. American Educational Research Association.
- Scott, S. S., McGuire, J. M., & Shaw, S. F. (2003). Universal Design for instruction: A new paradigm for adult instruction in postsecondary education. *Remedial & Special Education*, 24(6), 369–379. <https://doi.org/10.1177/07419325030240060801>.
- Silver, P., Bourke, A., & Strehorn, K. C. (1998). Universal Instructional Design in higher education: An approach for inclusion. *Equity & Excellence in Education*, 31(2), 47–51. <https://doi.org/10.1080/1066568980310206>.
- Snyder, T. D., de Bray, C., & Dillow, S. A. (2019). *Digest of education statistics 2017* (NCES 2018-070). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. <https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2018070>.
- Spirtes, P., Glymour, C. N., & Scheines, R. (2000). *Causation, prediction, and search*. MIT Press.
- Tobin, T. J., & Behling, K. (2018). *Reach everyone, teach everyone: Universal Design for learning in higher education*. West Virginia University.
- Tzivinikou, S. (2014). Universal Design for Learning - Application in higher education: A Greek paradigm. *Problems of Education in the 21st Century*, 60, 156–166.
- Wang, J., Dawson, K., Saunders, K., Ritzhaupt, A. D., Antonenko, P., Pasha, Lombardino, L., Keil, A., Agaclic-Dogan, N., Luo, W., Cheng, L., & Davis, R. O. (2018). Investigating the effects of modality and multimedia on the learning performance of college students with dyslexia. *Journal of Special Education Technology*, 33(3), 182–193. <https://doi.org/10.1177/0162643418754530>.
- Wellman, G. S., & Marcinkiewicz, H. (2004). Online learning and time-on-task: Impact of proctored vs. unproctored testing. *Journal of Asynchronous Learning Networks*, 8(4), 93–105.
- What Works Clearinghouse. (2020). *What works clearinghouse procedures handbook, version 4.1*. National Center for Education Evaluation and Regional Assistance. Institute of Education Sciences, U.S. Department of Education. <https://ies.ed.gov/ncee/wwc/handbooks>.
- Zhang, H., Nurius, P., Sefidgar, Y., Morris, M., Balasubramanian, S., Brown, J., Dey, A. K., Kuehn, K., Riskin, E., Xu, X., & Mankoff, J. (2021). How does COVID-19 impact students with disabilities/health concerns? *ArXiv:2005.05438 [Cs]*. <http://arxiv.org/abs/2005.05438>.