UDL Instruments and Measurements

Anya S. Evmenova  
George Mason University  
Fairfax, VA, USA  
aevmenov@gmu.edu

Kimberly M. Johnson  
Minnesota State University  
Monkato, MN, USA  
kimberly.johnson-1@mnsu.edu

Cynthia M. Okolo  
Michigan State University  
East Lansing, MI, USA  
okolo@msu.edu

Eric J. Moore  
University of Tennessee  
Knoxville, TN, USA  
Emoore50@utk.edu

ABSTRACT
This presentation will provide an overview of instruments and measures used in UDL educational environments. Based on a review of published UDL research, we will present a sample of existing instruments to measure UDL implementation. In addition, we will focus on approaches to measuring student outcomes in UDL settings that are related to being an “expert learner”: (a) purpose/motivation/work ethic; (b) resourcefulness; (c) executive functioning; (d) social skills; (e) basic academic skills; (f) content knowledge; and (g) engagement. Finally, challenges in UDL measurements will be discussed.

INTRODUCTION
Universal Design for Learning (UDL) is reaching more and more classrooms every day. Teachers and students are excited about this new framework of curriculum development that offers options for how content is presented; how students demonstrate what they have learned; and for engagement. However, few studies report data on both student outcomes and fidelity of UDL implementation. Thus, a workgroup was formed as part of the UDL-IRN community to review published research to categorize the types of instruments in use and to gauge the degree to which they are validated. A longer-term goal of the committee is to set up a network from which instruments can be shared, and possibly validated, among researchers.

UDL IMPLEMENTATION MEASURES
While educators are encouraged to use UDL in their classrooms to support students with diverse abilities and needs, there is no mutual understanding of what UDL implementation entails. UDL is a complex framework organized by three principles, nine guidelines, and 31 checkpoints. So how can we measure the implementation of UDL? In order to address this question, the workgroup has conducted a review of published UDL research. A review of 89 original research-based studies published in peer-reviewed journals in which the authors claimed to implement UDL as well as Universal Design of Instruction (UDI), or Universal Instructional Design (UID), was conducted. Based on the review, 36% of those articles included implementation measures. Moreover, only 22% of the reviewed articles included an instrument or a detailed description of the measures. Those implementation measures can be organized by several categories, which include educator perceptions and student attitudes, lesson plan reviews, and observation instruments.

Educator Perceptions and Student Attitudes
Published UDL research often reports teacher perceptions and student attitudes (including both K-12 and college students). A few studies used the revised versions of instruments validated by previous research such as Attitudes Toward Inclusion (Van Laarhoven, Munk, Lynch, Bosma, & Rouse, 2007) and Individualized Classroom Environment Questionnaire (ICEQ; Abell, Jung, & Taylor, 2011). In other words, the existing items might be modified to reflect the components of UDL. For example, compare the original item from ICEQ “Students choose their partners for group work” to the revised “Students choose the way they learn best to complete assignments (e.g., using technology, working independently, working with partners”). However, most of the studies in this category relied on researcher-developed questionnaires and interview protocols. These focused on the helpfulness and usefulness of various UDL-based materials and course elements from both the students’ (e.g., Black, Weinberg, & Brodwin, 2015; Kumar & Wideman, 2014; Rao, Edelen-Smith, & Wailehua, 2015) and educators’ perspective (e.g., Coy, Marino, & Serianni, 2014; LaRocco & Wilken, 2013) as well as the perceptions on the alignment to UDL guidelines (e.g., Scott, Temple, & Marshall, 2015). For example, “I like having short weekly assignments with a relatively low point value instead of fewer assignments that are worth more” (Rao & Tanners, 2011) or “Instructor presents information in multiple formats (Schelly, Davies, & Spooner, 2011). Those studies that focused on online or blended learning environments often also reported student usage statistics (e.g., Bongey, Cizadlo, & Kalnbach, 2010). Some studies examined the perceived preparedness of educators to implement UDL (e.g., Scott, Thoma, Puglia, Temple, & D’Aguilar, 2017; Vitelli, 2015).

Lesson Plan Reviews
All studies in this category used the original or modified rubric developed by Spooner et al. (2007) to evaluate the UDL principles included in lesson plans (e.g., Navarro, Zervas, Gesa, & Sampson, 2016). The rubric allows users to rate the use of representation, action/expression, and engagement components on a 3-point scale (0 = no clear
description of instructional modifications; 1 = one or two modifications discussion; 2 = three or more modifications discussed). In some studies, the use of the rubric was supplemented by the qualitative analysis of materials, instructional methods, and assessments proposed (e.g., Courey, Tappe, Siker, & LePage, 2012).

Observation Instruments.
Classroom observations were conducted in various learning environments (K-12, higher education, online). In most cases the instruments were researcher-developed and closely aligned with the purpose of the study. For example, the Inclusive Classroom Observation Tool (Morningstar & Shogren, 2013) was designed to include multiple indicators of effective classroom-inclusive practices such as UDL. The UDL Scan Tool was developed and validated to examine the alignment of online learning content to UDL framework (Basham, Smith, & Satter, 2016).

STUDENT OUTCOMES IN UDL SETTINGS
The second focus of our workgroup this year was to begin work related to identifying clearly defined theoretical outcomes of UDL and curating measures related to these outcomes. This is critical for advancing the work of researching UDL, because until we have a clear sense of the intended outcomes of UDL, any discussion around the claim that “UDL works!” may ring hollow. We identified three significant barriers related to measuring outcomes of UDL.

1. UDL is operationalized in different ways in research and practice. There may not be clarity or consensus as to what these outcomes are/should be.
2. Outcomes related to UDL may not be measurable in the static, traditional sense.
3. Measures may not exist for UDL outcomes, or may exist only through creation by UDL advocates, which may introduce potential for unintentional researcher bias.

We set out to semi-systematically address these barriers. To identify theoretical outcomes, we began by reviewing contemporary UDL texts such as Meyer, Rose and Gordon’s UDL Theory and Practice (2015). According to these authoritative authors, the primary focus of UDL is to develop expert learners; that is, learners who are in the perpetual process of becoming ever more resourceful, knowledgeable, strategic, goal directed, purposeful and motivated. Given both the authority of this text and its authors and the wide range of potentially measurable indicators that come from the construct of “expert learners” and its many sub-constructs, we decided to begin our work with this construct.

It is important to note that Meyer, Rose and Gordon caution against seeing these as endpoints which students could achieve, choosing to focus on such as being “always
a process of continuous learning–practice, adjustment, and refinement” (p. 21). Such dynamism, often viewed as a core strength of UDL, may be taken to pose challenges for measuring outcomes. The act of measurement requires taking samples at given moments, and thus introduces a degree of static that some may consider at odds with UDL theory. This fact collided with the second barrier we anticipated.

We suggest that this potential conflict can be resolved in the interpretation and reporting of findings from such measures. Indeed, it has been well argued in the literature (e.g., Gall, Gall & Borg, 2007) that it is not the measures themselves that are valid or invalid, but their interpretation that may or may not be valid. Valid interpretation must be guided by context and theory. In this case, instead of using measures of student performance in relation to resourcefulness and/or motivation and/or knowledge as endpoint indicators that students “have become” expert learners, a potentially more valid use and interpretation of such measures would be to employ them as pre- and post-tests to demonstrate growth across time in these domains. Such data may be used to implicitly indicate the presence of developing expertise. More discussion on this topic among the larger UDL community is warranted.

Finally, to address the third barrier, we drew from the sub-constructs of “expert learner” to identify several potentially measurable domains closely related. A sample of our current list of domains includes:

- Purpose/Motivation/Work Ethic
- Resourcefulness
- Executive functioning (e.g. Goal-directedness, strategic)
- Social skills (inc. intrapersonal such as self-concept, collaboration)
- Basic Academic Skills (e.g., Literacy)
- Content Knowledge
- STEM Specific Content Knowledge
- Communication Skills (oral, written, visual)
- Linguistic development (especially for ELL, students with disabilities affecting language acquisition).
- Engagement in class/instruction (could be measured by reduced behavioral issues, qualitative measures of engagement).

In a spreadsheet, we have begun a matrix with these domains and corresponding measures with information about each. At present, searches have been conducted via Mental Measures Yearbook, Web of Knowledge / Web of Science, ERIC including dissertations, and Psychosocial Instruments. We’ve compiled several measures thus far, though our matrix is far from comprehensive.

The next steps include continuing to fill out the matrix (potentially adding more domains as warranted), evaluating the appropriateness of measures included, and conducting UDL research with vetted measures. We invite the larger UDL community to contribute to the matrix, which may found at http://bit.ly/UDLMeasures (case sensitive).

Once we have a strong base of measures, our goal is to publish this matrix in a usable format that enables researchers or teachers to search for measures that suit their needs and the areas of expertise they are seeking to develop in their students in a given UDL environment or intervention.

CONCLUSION
As UDL gains more momentum, we need to do a better job at measuring UDL. As of now, the majority of UDL research is focused on measuring perceptions and attitudes. Moreover, the vast majority of measures do not meet the reasonable standards for technical adequacy. Instruments with greater validity are needed to offer both the practitioners and the researchers rigorous ways of measuring UDL implementation and student outcomes in UDL environments. The UDL Instruments and Measurements workgroup looks for input from the Summit attendees in their recommendations for future tasks.

REFERENCES
Kumar, K. L., & Wideman, M. (2014). Accessible by design: Applying UDL principles in a first year undergradu-


