REVIEW ARTICLE

Unraveling Challenges with the Implementation of Universal Design for Learning: A Systematic Literature Review

Ling Zhang[1] · Richard Allen Carter Jr[2] · Jeffrey A. Greene[1] ·
Matthew L. Bernacki[1]

Abstract

Educators and instructional designers have used the Universal Design for Learning (UDL) framework to guide their design of inclusive instruction for students with and without disabilities. Despite UDL having entered its 4th decade of development and research, there have been ongoing critiques of UDL for lacking clarity in definition, challenges with implementation, and insufficient evidence of its effectiveness. These critiques warrant further evaluation of UDL, especially with a focus on the theoretical underpinnings behind its conceptualization and implementation. Thus, we synthesized 32 peer-reviewed studies published between 1999 and 2023 that focused on UDL implementation in preK-12 educational settings and measured various aspects of student learning outcomes (e.g., cognitive, motivational, and behavioral). Specifically, we evaluated each study's intervention or instructional design in terms of its alignment to UDL checkpoints, guidelines, and/or principles as well as existing theories of learning or instructional design. Results revealed several interrelated challenges that stymie UDL research, including the absence of explicit alignment between UDL checkpoints and intervention or instructional designs investigated in the extant literature, the uneven coverage of implemented checkpoints and corresponding guidelines, the overlap among multiple checkpoints and guidelines, and the lack of theoretical guidance regarding the design and implementation processes. Based on these findings, we provide recommendations for strengthening the research base for less frequently applied UDL checkpoints, recommendations for documenting checkpoints and relationships among checkpoints as indispensable components of UDL implementation, and directions for future research conducted via systematic UDL implementation guided by established theories.

Keywords Universal design for learning · Inclusive instruction · Theoretical alignment · Literature review

---

Extended author information available on the last page of the article

Universal Design for Learning (UDL; Rose & Meyer, 2002), which is a framework for guiding the design of flexible educational practices, has entered its 4th decade of development and research. UDL is an extension of the *universal design* movement in architecture in the 1990s (e.g., King-Sears, 2009). The aim of universalization in architecture has been to make environments, services, and products accessible to and usable by the widest range of users, including individuals with disabilities, at the outset with no or minimal need for retrofitting, adaptation, or specialized design (Mace, 1997). Scholars at CAST, which is the originator of UDL, applied the concept of *universal design* to education (Rose & Meyer, 2002). UDL highlights the necessity of building flexibility in the curriculum (i.e., goals, materials, methods, and assessments as defined by CAST) to support and improve access to equitable learning opportunities for all students with and without disabilities (Meyer et al., 2014). Through the UDL lens, any curriculum that is not designed proactively to account for learner variability is "disabling" because it fails to address the diverse needs of individual learners (Rose & Meyer, 2002). The ideology behind UDL centralizes "fixing" curriculum rather than remediating student deficits (Rose & Meyer, 2002; Waitoller & Thorius, 2016). Given the UDL approach to designing curriculum is complex, educators and researchers need tools and guides to inform their development of "enabling" curricula, instruction, and interventions.

To turn the concept of universal design into an actionable educational construct, researchers created the UDL framework and articulated it through a graphic organizer called the UDL Guidelines (see version 2.2 at https://udlguidelines.cast.org/). The UDL Guidelines consist of three overarching principles which include providing (1) multiple means of representation, (2) multiple means of action and expression, and (3) multiple means of engagement, which are broken down into nine guidelines (CAST, 2018a). Each guideline is further illustrated through two to five checkpoints that provide suggestions for designing learning experiences tailored to individual students (Rao et al., 2014). Educators, instructional designers, and other stakeholders can use the framework to systematically anticipate and reduce barriers to learning via implementing the principles and guidelines (Ok et al., 2017). Supporting students with multiple pathways and tools to learn and achieve learning goals (e.g., cognitive, motivational, and behavioral outcomes) is central to UDL conceptualization and the desired outcome of its implementation (Rose & Meyer, 2002).

Over the past three decades, there has been increasing enthusiasm for the UDL framework among researchers, policymakers, and practitioners as evidenced by an exponential increase in articles about the framework and prevalent references to the framework in educational policies (Hollingshead et al., 2022). Research on UDL across contexts is accumulating, including applications in special and general education, preK-12 and higher education, as well as professional learning in education and industry (Capp, 2017; Fornauf & Erickson, 2020; Kennedy et al., 2018; Ok et al., 2017). On the policy side, UDL is referenced in multiple US federal education policies, such as the Higher Education Opportunity Act of 2008, 2008 and the Every Student Succeeds Act of 2015, 2015.

Along with the enthusiasm for UDL have also come critiques regarding both the conceptualization and the implementation of the framework. An ongoing critique is that UDL lacks theoretical underpinnings, operationalization, and empirical evidence (e.g.,

Boysen, 2021). As this critique has increasingly manifested in conceptual and empirical articles (e.g., Webb & Hoover, 2015), addressing this critique is essential for UDL research to move forward. More importantly, there is a pressing need to conceptualize, validate, and ground the implementation of UDL in existing theories of human learning and/or instructional design. Researchers have long called for addressing the ambiguity of defining, implementing, and measuring the effects of the UDL framework (e.g., Kennedy et al., 2014). In one critique, Edyburn (2010) argued the conceptualization and measurement of UDL implementation have not been clear and indicated "the claim that UDL has been scientifically validated through research cannot be substantiated at this time" (p. 34). Then, he proposed new directions for researchers to generate adequate research evidence over the next decade of UDL scholarship. For example, some directions included situating UDL in the field of instructional design, clearly defining UDL, and measuring the contributions of UDL to the development of expertise.

In a follow-up article 10 years later, Edyburn (2021) stressed four ongoing challenges with UDL, including "(a) definitional clarity about a UDL intervention; (b) inability to isolate the active ingredients thought to make UDL effective; (c) guidelines about the dosage of UDL intervention need to achieve access, engagement, and success; and (d) appropriate research methodologies relevant to the standard of evidence-based practice" (p. 308). These critiques indicate the lack of progress in conceptualizing, defining, implementing, and measuring the effects of the UDL framework. Such progress seems possible, given that UDL scholars have suggested that new advances in learning sciences, neurosciences, and other related disciplines could inform the research investigating UDL as an evolving construct (e.g., Rappolt-Schlichtmann et al., 2012). Nonetheless, currently, there has been little effort to examine UDL or its implementation from the perspective of learning or instructional design theories.

These ongoing critiques of the UDL framework warrant further evaluation of the framework as it enters its 4th decade, with a particular focus on the theoretical underpinnings behind its conceptualization and implementation. In this article, we reviewed the extant literature to investigate why Edyburn's (2021) challenges still beguile UDL implementation and what efforts are needed to make progress on these challenges. First, we discussed the conceptualization behind UDL, issues that have led to the ongoing UDL critiques, and gaps in previous literature reviews and meta-analyses of UDL research to address identified issues. Then, we systematically analyzed research that investigated UDL implementation in preK-12 settings. To address gaps in previous reviews, we focused on analyzing how researchers applied UDL principles, guidelines, and checkpoints as well as evaluated the extent to which UDL implementation aligned to learning or instructional design theories. Last, we provided implications for future UDL research.

Conceptualization of the UDL Framework and Associated Challenges

According to the early UDL creators, Rose and Meyer (2002), the development of the UDL framework was grounded in research from neuroscience, learning sciences, and other areas. They indicated the three-principle framework of UDL was

predicated upon neuroscientific research that substantiated variability among individual learners with regard to three learning-related neural networks (i.e., recognition, strategic, and affective). Rose and Meyer (2002) associated the recognition networks with the posterior cortex, the strategic networks to the frontal cortex, and the affective networks to the medial cortex. The concept behind UDL was to anticipate and address learner variability by providing flexible options for learners to engage in learning (i.e., aligned to affective networks), access information (i.e., aligned to recognition networks), and express understandings of knowledge and skills (i.e., aligned to strategic networks; Meyer et al., 2014). Later, UDL scholars (e.g., Rappolt-Schlichtmann et al., 2012) acknowledged that the neurological organization of UDL oversimplified the complexity found in the brain sciences.

Frequently, researchers and policymakers have described UDL as a "scientifically validated" framework by referring to empirical evidence for each individual UDL guideline and checkpoint (e.g., HEOA, 2008; Hollingshead et al., 2022). However, this raises a critical question, "Does research on individual guidelines and checkpoints, which was conducted outside the UDL context, demonstrate the effectiveness of UDL as a design framework?" We argue the answer is no because research is needed on the efficacy of UDL principles, guidelines, and checkpoints as a coherent design framework. Similarly, UDL scholars have called for more rigorous research to evaluate UDL implementation (e.g., Basham et al., 2020; Smith et al., 2019).

The lack of research targeting the UDL framework itself also means researchers have not yet reached a consensus on essential elements or "active ingredients" that constitute rigorous UDL implementation (Edyburn, 2021). As UDL guidelines and checkpoints encompass a wide range of research topics and can be applied in a myriad of ways, Hollingshead et al. (2022) raised a concern that "Without that [a shared understanding of how to define and measure UDL], everything is UDL and UDL is nothing" (p. 1137). In an attempt to clarify the ambiguity of UDL, Hollingshead et al. (2022) examined how experts in the development of and research on UDL defined the framework and its critical features. Rather than offering clarity, Hollingshead and colleagues found substantive differences and continued discrepancies regarding UDL definitions and implementation among these experts. The ongoing use of ambiguous definitions and justifications for UDL, and the subsequent implementation challenges that result, remain the biggest hurdles for UDL research.

In an attempt to establish a collective research agenda on UDL, a committee of scholars and practitioners from the Universal Design for Learning–Implementation and Research Network (UDL-IRN) indicated that "UDL is not simply a listing of various flexible options and strategies;" instead, it is a *systematic* design and implementation process (Smith et al., 2019, p. 177). In their statement, the committee recommended that researching systems of UDL practices would contribute to clarifying the ambiguity of implementing UDL as a coherent framework; however, they did not further define what "systematic" or "systems" means for UDL implementation. Clearly defining approaches to systematically implementing UDL will help the field move beyond the current stagnation in UDL research.

Implementing a complex system in a theory-aligned way can help enhance the conceptual clarity of the system and identify essential implementation elements (Nilsen, 2015). We argue that theory-guided perspectives provide critical insights

into the *systematic* implementation of UDL as a complex framework, which echoes Rappolt-Schlichtmann et al.'s (2012) call for utilizing advances in learning sciences to inform UDL research. This perspective requires a closer examination of educational theories that informed the development of UDL. The early developers of UDL applied a variety of learning and instructional design theories to develop checkpoints and guidelines, for example, Mayer's (2005) cognitive theory of multimedia learning (CTML) as well as Deci and Ryan's (2002) Self-Determination Theory (SDT). To provide a comprehensive picture, we analyzed and identified theories that were explicitly referenced on the CAST website as theoretical underpinnings for each guideline and checkpoint (see Supplemental Table 1). Despite the large number of referenced theories, questions remain as to how well researchers implemented UDL guided by these theories, further warranting more research to investigate theoretically aligned approaches, models, or systems of UDL practices.

Previous Reviews of UDL Research

There are many review studies that synthesized characteristics of UDL implementation and evaluated the effectiveness of UDL-based interventions for different populations, such as preK-12, postsecondary, and rehabilitation health professionals (e.g., Capp, 2017; Ok et al., 2017). Most literature reviews reported whether primary studies explicitly aligned their interventions or instruction to UDL. A common finding shared by most reviews is that researchers applied UDL guidelines and checkpoints in varied combinations across different contexts (Rao et al., 2014). However, few reviews provided detailed descriptions of how interventions or instruction in primary studies align to UDL principles, guidelines, or checkpoints nor assessed the quality of these studies (e.g., Al-Azawei et al., 2016; Capp, 2017).

Regarding the effectiveness of UDL, a recent meta-analysis conducted by King-Sears et al. (2023) found a moderate positive effect for learners receiving UDL-based interventions ($g = 0.43$), with a stronger impact for school-aged learners ($g = 0.48$) than adult learners ($g = 0.28$). According to King-Sears et al. (2023), their study was the first and only methodologically-sound meta-analysis that analyzed the effect of UDL-based interventions. Other researchers claimed to have conducted a meta-analysis (e.g., Al-Azawei et al., 2016; Capp, 2017) but reported no quantitative meta-analysis of effect sizes of UDL-based interventions, thus affirming that the King-Sears et al. (2023) study was the only meta-analysis of research investigating the effects of UDL-based interventions on students' academic achievement in content areas.

Additionally, King-Sears et al. (2023) assessed the quality of primary studies on UDL using the UDL Reporting Criteria (UDL-RC). The UDL-RC is a rubric that was recently developed to establish quality indicators for reporting UDL-aligned design and implementation (Rao et al., 2020). This rubric guides researchers in articulating three criteria, including descriptions of (a) Learner Variability and Environment, (b) Proactive and Intentional Design, and (c) Implementation and Outcomes (Rao et al., 2020). Accordingly, King-Sears et al. (2023) assessed whether primary studies included in their meta-analysis provided these descriptions of UDL

implementation (i.e., yes or no). The researchers summarized that 45% of studies (*n* = 9) provided detailed descriptions of how nine UDL guidelines and 31 checkpoints were applied. Like other reviews, however, there were no in-depth descriptions of how and which of the 31 checkpoints were applied to the intervention in a study-by-study format. Thus, it remains unclear how well researchers applied the UDL framework to the designs of interventions or instruction in empirical studies. Additionally, no previous reviews have analyzed whether UDL was implemented in a theoretically guided way. These gaps indicate the need to further dissect UDL implementation and evaluate its alignment with insights from learning sciences and theories.

The Present Review

To address the ongoing critiques, calls from the field, and gaps in the previous review studies, we conducted this literature review to investigate how researchers utilized UDL checkpoints, guidelines, and principles with a focus on identifying patterns and offering clarity in UDL implementation. Additionally, we evaluated if the design and implementation of UDL-based interventions or instruction were guided by learning or instructional design theories. We sought to examine the existing empirical support for UDL implementation in preK-12 settings and determine how well it aligned with UDL itself as well as the theories that informed its development. As such, we focused on studies that reported outcomes of student learning and/or its various aspects, such as cognitive, metacognitive, motivational, affective, engagement, and social-emotional learning outcomes. The following research questions (RQs) guided our inquiry:

1. How did UDL implementation align with UDL checkpoints, guidelines, or principles, and what patterns or gaps emerged from the extant literature?
2. Whether and how was UDL implementation in preK-12 educational settings guided by learning or instructional design theories?

Method

We employed the theories-characteristics-contexts-methods (TCCM) framework (Paul et al., 2023) to guide the synthesis of previous studies on UDL implementation in preK-12 educational settings. The TCCM framework provides an analytic tool for structuring thematic reviews of literature and suggesting future research directions regarding theories, characteristics, contexts, and methods (Paul et al., 2023). Guided by the TCCM framework, we identified and organized pertinent information on the theoretical underpinnings of UDL implementation, contexts shaping research settings, characteristics of UDL implementation, and methodological approaches. The first and second authors, who previously conducted research on UDL implementation in preK-12 settings, conducted the literature search and coding.

Literature Search

Following the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guidelines (Page et al., 2021), we searched empirical studies that reported UDL implementation and its effects on different types of student learning outcomes in preK-12 settings (see Fig. 1). In this study, we defined student learning outcomes as including academic achievement in subject content areas, metacognitive skills (e.g., self-regulated learning skills; Greene, 2017), affective-motivational outcomes (e.g., interest), behavioral functioning (e.g., increase of expected behaviors), and social-emotional skills (e.g., social skills). We used four databases to search peer-reviewed articles published between 1999 and May 4, 2023 (the search date). These databases include Academic Complete Search Premier, ERIC, PyscINFO, and MedPub, which represent the most widely used databases in education, psychology and behavioral science, and medicine, respectively. The starting year was chosen because it was when CAST started to disseminate UDL through education policies and reports (Edyburn, 2010).

We conducted the search using boolean terms in the title and abstract: ("Universal Design for Learning" OR "UDL" OR "Universal Design") AND ("outcome" OR "impact" OR "effect" OR "affect" OR "emotion" OR "cogniti*" OR "metacogniti" OR "learning" OR "skill" OR "knowledge" OR "motivat*" OR "engagement" OR "behavio*r" OR "emotion" OR "enjoyment" OR "performance" OR "academic"). We delimited our search to documents written in English and published in peer-reviewed journals. The search yielded 2633 documents. After removing duplicates, 1548 records were uploaded to Rayyan, an artificial intelligence-powered software using a semi-automation process to expedite screening processes for systematic reviews (Ouzzani et al., 2016). We used the following inclusion and exclusion criteria to identify studies.

Inclusion and Exclusion Criteria

First, to control for variation in student populations, studies included in this review reported on UDL implementation in preK-12 settings; thus, we excluded studies focused on higher education (see a review in Fornauf & Erickson, 2020), rehabilitation (see a review of UDL implementation by rehabilitation health professionals in Kennedy et al., 2018), or corporate settings (e.g., Irbe, 2016). We applied this criterion to identify gaps in UDL research specifically for the preK-12 student population.

Second, we included studies that evaluated different types of student learning outcomes (i.e., academic, metacognitive, affective, motivational, behavioral, and social-emotional). For example, studies were included if one or more student academic or metacognitive outcomes were assessed through measures, such as standardized assessments or researcher-developed performance measures. Studies were included if student affects, interests, engagement, or motivation toward their academic learning were assessed, such as by student self-reported and teacher-reported surveys. Studies were included if student social skills or behaviors, including behaviors in
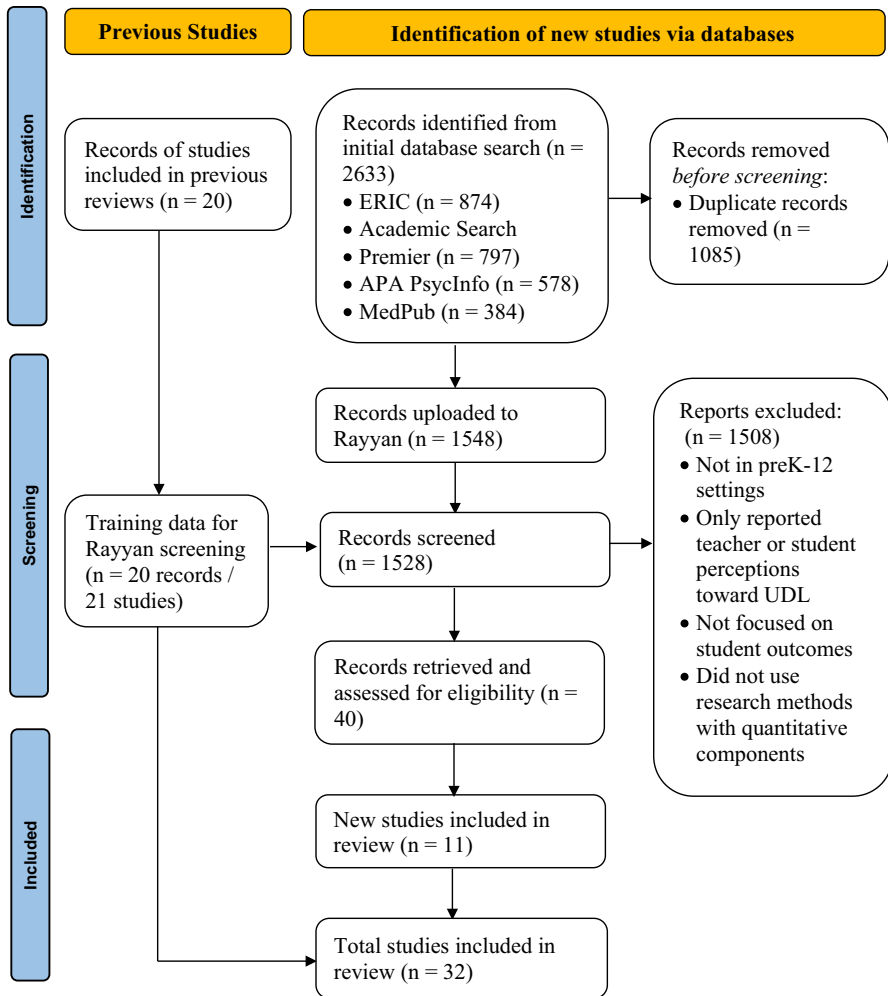
Fig. 1 PRISMA literature search and screening procedure

online learning, were assessed through surveys or other observation data (e.g., digital trace data; Bernacki, 2018). We excluded studies that solely reported teacher or student perceptions of characteristics of UDL-based interventions or environments, rather than perceived changes in student learning outcomes resulting from participating in UDL-based interventions or environments. For example, we excluded the Abell et al. (2011) study because researchers only measured student perceptions of opportunities to interact with teachers, participate in learning, or make decisions within UDL-aligned environments, rather than evaluating perceived changes in academic performance or engagement due to participation in UDL-aligned instruction. On the other hand, we included the Kortering et al. (2008) study because researchers evaluated students' perceived engagement after participating in a UDL intervention.

Additionally, we excluded studies that only focused on professional learning for pre- and in-service teachers, such as training pre-service teachers in using UDL in lesson planning (e.g., Courey et al., 2013).

Third, we included only empirical studies with randomized controlled trials (RCT), quasi-experimental designs (QED), mixed-method research (MMR), single-case design (SCD), regression discontinuity designs (RDD), or other methods with quantitative components. Additionally, we excluded dissertations and conference proceedings to avoid potential duplicate publications. It is important to note that the goal of this review was not to calculate effect sizes of UDL-based interventions given that the recent meta-analysis reported in King-Sears et al. (2023) provided up-to-date information on the effects of UDL-based interventions in preK-12 settings. Instead, we focused on dissecting information on UDL design and implementation.

Identification of Relevant Studies

We used two steps to identify articles eligible for inclusion. First, we used our inclusion and exclusion criteria to identify 21 studies with 20 articles, sourced from ten previous review studies (e.g., Beerwart, 2018; King-Sears et al., 2023). These studies involved empirical research on UDL-based interventions that met our inclusion criteria. Next, we used these articles as training data for the Rayyan software. Rayyan uses a support vector machine classifier to learn from users' labeling of studies to be included or excluded and then outputs a score of how close the remaining non-labeled studies match the included studies (Ouzzani et al., 2016). Using the training data, Rayyan generated a ranking of the remaining studies ($n = 1528$) for their relevancy to our inclusion criteria and research questions. Second, the first and second authors screened titles/abstracts of the remaining records independently using the relevancy ranking generated by Rayyan to eliminate clearly ineligible articles, such as non-UDL interventions.

After the independent screening process, the two raters agreed on excluding 93.2% and including 6.4% of articles for further examining eligibility as well as disagreed on labeling 0.4% of articles. This process yielded 6.8% of articles ($n = 104$) for further determination. Then, the two raters screened the titles/abstracts of 104 articles together and retrieved full texts of 40 records to examine against the inclusion and exclusion criteria for eligibility for inclusion. This process of determining eligibility of the 40 full texts yielded another ten studies that met our inclusion criteria. However, we excluded Marino et al. (2010) from the analysis given that it used the same dataset as that of Marino et al. (2010). In total, we included 32 studies from 31 articles in the present review, including two studies by King-Sears and Jackson (2020). We calculated the inter-rater reliability (IRR) for the process of screening and identifying relevant studies using the formula agreements/(agreements + disagreements) × 100. An IRR of 93% was achieved during this phase. The two coders convened to discuss articles that did not have consensus and resolved all discrepancies by determining the eligibility of these articles against inclusion criteria.

Coding Scheme and Procedure

We created a four-part coding scheme by adopting and integrating the TCCM framework, Methodological Quality Indicators (MQIs; Miller et al., 2018), and UDL-RC (Rao et al., 2020). The TCCM framework guided us in categorizing the codes into four areas regarding UDL implementation, including theory, context, characteristics, and methodology (see Supplemental Table 2). Specifically, MQIs consisted of seven quality indicators for evaluating each included study regarding its theoretical alignment and research base, rigor of methodology, and consistency of the study findings (Miller et al., 2018). As indicated above, UDL-RC was a rubric consisting of eight criteria within three categories that could be used to assess critical features of UDL design and implementation. The rubric served as a non-evaluative tool that enables researchers to identify whether an element is present or not (i.e., using "yes" or "no" to rate each criterion; Rao et al., 2020). All authors agreed upon the coding scheme before the two coders began the coding process.

Coding for Theory and Methodology

In alignment with the *theory* and *methodology* of TCCM, we used two indicators under MQI Standard 1—Theoretical Alignment to identify (a) whether UDL implementation aligned to theory, and if yes, what theory was applied; and (b) whether findings were linked to theory. Given that theory and previous research were both evaluated in this standard, we coded included studies by specifying whether they linked arguments to "theory only," "research only," "both," or "neither." For studies that explicitly linked arguments to theories, we further evaluated each study in terms of alignment between UDL implementation and theoretical guidance. Specifically, we coded (a) whether the author(s) provided justifications for designing intervention or instruction guided by theories; (b) how the theoretical guidance informed the selection of UDL principles, guidelines, and/or checkpoints (i.e., the degree of theoretical alignment); and (c) whether findings on student learning outcomes substantiated the theoretically guided designs.

The other five MQI indicators were applied to each study to score their methodological quality. Each indicator was rated by using "yes" or "no" to indicate if a study met the criterion or not; correspondingly, a 0–7 scoring system was applied to determine the methodological quality of included studies. Moreover, to extract more information, we generated codes that enable us to evaluate each study for the specific method applied (e.g., experimental design, single-case design), student learning outcomes (e.g., academic, motivational, and behavioral), evidence categories (e.g., positive causal relation, positive correlational, no/mixed evidence, negative causal, and negative correlational), and fidelity data.

Coding for Context and Characteristics

To gather information about *contexts* and *characteristics* specifically for UDL implementation, codes were developed in alignment with three UDL-RC criteria. For the *contexts*, we adopted UDL-RC Criterion 1 Learner Variability and Environment to

code student demographics (i.e., age, race/ethnicity, and disability category) and learning environments (i.e., general, inclusive, or special education settings and content areas) reported in included studies. For the *characteristics*, we evaluated the extent to which UDL implementation was described in each study, which was in line with UDL-RC Criterion 2c Application of UDL Guidelines and Checkpoints as well as 3a Description of Implementation of Practice/Intervention.

There were studies that *explicitly referenced* UDL principles, guidelines, and checkpoints as design components of the intervention or instructional practices using checkpoint/guideline numbers (e.g., guideline 1, checkpoint 1.1) or language linked to the framework (e.g., multiple means of representation [principle], options for perceptions [guideline 1], and customizing the display of information [checkpoint 1.1]). For example, King-Sears and Johnson (2020) used the language "options for mathematical expressions (support decoding of mathematical notations)" when describing their intervention. The language explicitly aligned to UDL guideline 2 "provide options for language and symbols" and checkpoint 2.3 "support decoding of text, mathematical notation, and symbols." Moreover, there were studies that did not specify checkpoint/guideline numbers nor use terminology signifying specific checkpoints, guidelines, or principles. For example, Coyne et al. (2012) described multiple design features of their intervention using such language as "sentence-by-sentence human digitized voice with synchronized highlighting." The researchers indicated that these designs aligned with the principle of multiple means of representation; however, they did not further specify checkpoints or guidelines.

The examples above suggested varying levels of specificity regarding how the researcher(s) of each study referenced UDL principles, guidelines, or checkpoints when describing the designs of their intervention or instruction. To code the level of specificity, we created four categories, including (a) explicit references to checkpoints and associated guidelines/principles, (b) explicit references to principles and associated guidelines but *not* checkpoints, (c) explicit references to principles but *not* guidelines *or* checkpoints, and (d) without explicit references to principles, guidelines, or checkpoints. We coded each study based on this categorization. Next, we assessed how the design features of intervention or instruction reported in each study aligned to UDL checkpoints, another detail that had not been described sufficiently in previous reviews (see Supplemental Table 3 for coded designs of each study).

For studies *with explicit references to checkpoints* (category a), we extracted specified checkpoint numbers or language. For studies *without explicit references to checkpoints* (categories b, c, and d), we assessed the alignment between intervention or instructional designs and UDL checkpoints following a deductive coding procedure (Saldaña, 2020). First, we used the 31 UDL checkpoints as a priori codes. Then, we extracted descriptions of the design features of the instruction or intervention, which were often reported in the method or other related sections of these studies. Next, we applied these codes to the excerpts where the described designs aligned with UDL checkpoints. To assess alignment, we consulted descriptions and implementation suggestions for each checkpoint outlined on the CAST website. For example, the above-mentioned design in Coyne et al. (2012) aligned with CAST's suggestion for "using digital text with an accompanying human voice recording,"

as detailed in checkpoint 2.3 under guideline 2 (see specific descriptions at https://tinyurl.com/UDLCheckpoint). We acknowledge that in studies without explicit references, researchers still applied UDL to design their interventions or instructions despite the absence of explicit language specifying checkpoints, guidelines, or principles. Therefore, by assessing the alignment of the design features reported in these studies to UDL checkpoints through our coding procedure, we ensured a comprehensive evaluation of UDL-guided designs with varying levels of specificity.

The two coders conducted the deductive coding jointly. Other aspects of the included studies were coded by the two coders independently. IRR was calculated by using Cohen's kappa coefficient (κ) with 95% confidence intervals using the formula: $\kappa = (Po - Pe) / (1 - Pe)$, where Po represents observed agreement, and Pe represents chance agreement (McHugh, 2012). The calculation yielded a $\kappa = 0.71$ for theory, $\kappa = 0.73$ for context, $\kappa = 0.87$ for characteristics, and $\kappa = 0.86$ for methodology; all exceeded the level of substantial agreement between coders (Landis & Koch, 1977). The two coders convened to resolve all discrepancies through open discussion where both coders first reviewed their own codes and then listened to the other's rationale for choosing a code, until achieving consensus.

Results

We analyzed 32 studies that investigated the implementation of UDL in preK-12 educational settings and evaluated its impact on student learning (see Supplemental Table 4 for detailed information on and full citations for each study). Table 1 summarizes the contexts of UDL implementation across synthesized studies. Overall, ratings for study quality ranged from 4 to 7 across all studies, indicating moderate to high quality as assessed through MQI. Most studies described students' disability status and/or provided a breakdown of student race/ethnicity backgrounds. For studies without sufficient information on learner characteristics, it was difficult to determine whether and how a UDL-based intervention or instruction benefited different groups of students with specific disabilities or from diverse cultural backgrounds. Additionally, research using experimental designs (i.e., RCT [$n = 6$] and QED [$n = 10$]) or SCD ($n = 4$) to investigate effects on student learning outcomes were scarce considering the wide span of years of UDL research and that such research was scattered across all preK-12 grade levels and multiple content areas (see Supplemental Table 5).

RQ1: Characteristics of UDL Implementation

References to Checkpoints, Guidelines, or Principles

Table 2 illustrates how UDL principles, guidelines, and/or checkpoints were applied in the extant literature. Only a few studies ($n = 8$; 25%; marked with a solid circle) explicitly referenced checkpoints in the descriptions of designs of UDL-based interventions or instructional practices. In five out of the eight studies, researchers

Table 1 Context of included studies ($n = 32$)

Category	Subcategory	Number	Percentage
Sample size	< 10	6	18.75%
	10 to 100	11	34.38%
	101 to 1000	13	40.63%
	> 1000	2	6.25%
Research method	Randomized controlled trial (RCT)	6	18.75%
	Quasi-experimental design (QED)	10	31.25%
	Single-case design (SCD)	4	12.50%
	Mixed method	8	25.00%
	Other (i.e., design-based, correlational, survey study)	4	12.50%
Grade level	PreK	2	6.25%
	Grades K–5	13	40.63%
	Grades 6–8	10	31.25%
	Grades 9–12	9	28.13%
	K–12	2	6.25%
Content area	Science	10	31.25%
	English Language Arts (ELA)	7	21.88%
	Mathematics	2	6.25%
	Social studies	2	6.25%
	Multiple areas (e.g., math, ELA, transition goal)	4	12.50%
	Behavior intervention	1	3.13%
	Physical education	1	3.13%
	Other (i.e., student engagement, test system)	5	15.63%
Student population	Only students without disabilities	4	12.50%
	Only students with disabilities	9	28.13%
	Both populations	19	59.38%
Disability status	Specific learning disabilities	12	37.50%
	Intellectual disabilities	8	25.00%
	Autism spectrum disorder	6	18.75%
	Other health impairments (e.g., ADHD)	6	18.75%
	Speech or language impairment	5	15.63%
	Emotional disturbance	3	9.38%
	Mental health diagnoses; multiple disabilities; hearing impairment; orthopedic impairment	2 (each)	6.25%
	Visual impairment	1	3.13%
	Students with disabilities, IEPs, or 504 plans	9	28.13%
Race/ethnicity	White/Caucasian	18	56.25%
	Black/African American	17	53.13%
	Hispanic/Latino	16	50.00%
	Asian	9	28.13%

Table 1 (continued)

Category	Subcategory	Number	Percentage
	American Indian or Alaska Native	5	15.63%
	Multiple races	6	18.75%
	Native Hawaiian or other Pacific Islanders	4	12.50%
	Indigenous learners (Canada)	2	6.25%
	Minority students	1	3.13%
	Not reported	10	31.25%

Studies were coded multiple times when information provided on grade level, race/ethnicity, or disability fell under multiple categories

specified both corresponding guidelines and principles for the referenced checkpoints. In one of the remaining three studies, Zhang et al. (2021) explicitly referenced a checkpoint and corresponding guideline. The other two studies (i.e., Marino et al., 2014; Hitchcock et al., 2016) explicitly referenced checkpoints using respective numbers even though they did not specify corresponding guidelines or principles.

In several studies with explicit references to checkpoints, there were instances where one design feature of intervention or instruction was aligned to two or more UDL checkpoints, guidelines, or principles. For example, Root et al. (2020) described a self-monitoring strategy implemented in their intervention as aligned to both checkpoints 6.4 and 9.3, corresponding to the principles of Action and Expression and Engagement, respectively. Another example of one design feature aligned to multiple checkpoints, guidelines, or principles was a self-management mnemonic strategy reported by King-Sears and Johnson (2020). This strategy was described in alignment with UDL's principle of Representation because it provided options for decoding mathematical notations. It also "corresponded to UDL's principle of *expression*, in which distractions were minimized (i.e., focus only on these steps), and self-regulation, which facilitated students' execution of the strategy" (King-Sears & Johnson, 2020, p. 210). However, designs associated with minimizing distractions (checkpoint 7.3) and self-regulation (guideline 9) are usually considered under UDL's principle of Engagement.

In most studies (*n* = 24; 75%), researchers did not specify checkpoints or guidelines when describing their intervention or instructional designs. Of these studies, 11 (marked with a half-solid circle) indicated that the intervention or instruction was designed in line with UDL principles. However, they did not specify checkpoints or guideline numbers; nor did they incorporate language with clear linkages to checkpoints/guidelines. For example, Coyne et al. (2012) examined universally designed e-books and indicated that various features of these e-books were designed based on UDL principles and grouped these features under each principle; however, they did not further specify guidelines or checkpoints to which these features aligned.

The other 13 studies (marked with a blank circle) broadly stated that their interventions or instructional designs were based on UDL; however, researchers did not specify any checkpoints, guidelines, or principles for the designs (e.g.,

Table 2 Alignment of the universal design for learning (UDL) interventions to checkpoints, guidelines, and principles

Study, alignment illustration, and number of aligned checkpoints			Representation — Language & Symbols			Comprehension					Comprehension				Action & Expression — Physical Action		Expression & Communication				Executive Functions			Engagement — Recruiting Interest			Sustaining Effort & Persistence				Self-Regulation		
			1.1	1.2	1.3	2.1	2.2	2.3	2.4	2.5	3.1	3.2	3.3	3.4	4.1	4.2	5.1	5.2	5.3	6.1	6.2	6.3	6.4	7.1	7.2	7.3	8.1	8.2	8.3	8.4	9.1	9.2	9.3
Akhigbe & Adeyemi, 2020	◐	9						x			x				x						x		x		x			x	x	x			
Browder et al., 2008	◐	5		x	x										x											x				x			
Coyne et al., 2012	◐	11			x	x		x	x			x	x		x			x			x			x									x
Craig et al., 2022	○	0																															
Daley et al., 2016	◐	9			x	x		x		x			x		x							x								x			x
Daley et al., 2020	◐	11			x	x		x	x						x		x	x					x	x	x					x			
Dalton et al. 2011	◐	15	x		x	x		x	x	x		x			x			x	x		x			x	x			x		x			
Dolan et al., 2005	○	2		x		x																											
Hall et al., 2015	●	13	x		x	x		x		x		x			x		x	x	x		x				x				x				
Hitchcock et al., 2016	●	9						x			x							x	x			x		x	x						x	x	
Katz, 2013	○	0																															
Katz et al., 2021	○	0																															
Kennedy et al., 2014	◐	2		x					x																								
King-Sears & Johnson, 2020 (1)	●	14		x	x	x		x		x		x	x					x	x		x		x	x		x							x
King-Sears & Johnson, 2020 (2)	●	14		x	x	x		x		x		x	x					x	x		x		x	x		x							x
King-Sears et al., 2015	●	13		x	x	x		x		x		x	x					x	x		x		x			x							x
Kortering et al., 2008	○	0																															
Lieber et al., 2008	◐	4								x								x	x						x								
Loman et al., 2018	◐	5		x	x										x									x	x								
Marino, 2009	○	5						x	x									x			x			x									
Marino et al., 2014	●	21	x	x	x	x		x	x	x	x				x	x	x	x	x	x	x		x	x				x	x	x	x		x
McMahon et al., 2016	●	4			x														x					x						x			
Proctor et al., 2011	◐	9			x				x	x		x	x		x			x						x						x			
Rappolt-Schlichtmann et al., 2013	○	8		x	x		x					x			x			x						x						x			x
Root et al., 2020	●	20	x		x	x	x				x	x		x	x		x	x	x	x	x		x	x	x	x	x	x		x			x
Scott et al., 2011	◐	4		x	x										x																x		
Sokal & Katz, 2015	○	0																															
Taunton et al., 2017	○	9		x	x										x			x			x			x	x					x	x			
Twyman & Tindal, 2006	○	3		x	x			x																									
Yavuzarslan & Arslan, 2020	○	0																															
Yu et al., 2021	○	9		x	x			x	x						x	x	x				x								x				
Zhang et al., 2021	●	1																															x
Total N of aligned checkpoints			4	9	**19**	11	4	13	4	14	4	10	8	**1**	11	**1**	8	12	9	4	**13**	6	4	**13**	8	5	3	4	5	11	**2**	3	5

Note. The symbol ○ indicates that the study did not explicitly reference UDL principles, guidelines, or checkpoints when describing the interventions or instructional designs. The symbol ◐ indicates that the study explicitly referenced UDL principles but not guidelines or checkpoints. The symbol ● indicates that the study explicitly referenced UDL checkpoints, checkpoints with guidelines, checkpoints with principles, or checkpoints with guidelines and principles.

Rappolt-Schlichtmann et al., 2012; Yu et al., 2021). Three studies (i.e., Sokal & Katz, 2015; Katz, 2013; Katz, 2021) investigated the effects of a three-block model of UDL that emphasizes socio-emotional learning, inclusive pedagogy, and systemic structures for supporting these processes. According to researchers, this model expanded "traditional UDL foci on technology and differentiation to explore both the social and academic practices of the classroom" (Katz, 2013, p. 4). Although building upon CAST researchers' conceptualization of UDL, this model does not explicitly incorporate the commonly used UDL principles, guidelines, or checkpoints. As another example, Craig et al. (2022) focused on examining the correlation between students' standardized test performance and district-wide UDL implementation rather than detailing features of UDL implementation; thus, the researcher did not describe specific UDL-aligned instructional designs. Similarly, Kortering et al.

(2008) broadly described 24 teacher-created UDL practices and indicated that the use of UDL practices in instruction was voluntary, leaving specific UDL-aligned instructional designs unclear. Lastly, Yavuzarslan and Arslan (2020) reported the effects of a course designed based on UDL but failed to detail the specific course designs.

The lack of explicit alignment led to confusion regarding how researchers applied UDL checkpoints, guidelines, and principles to inform their intervention or instruction. This challenge emerged when we coded and aligned the design features of intervention or instruction to specific UDL checkpoints for studies without explicit references to checkpoints. For example, Hall et al. (2015) examined a UDL-based digital reading environment that provided access to multiple features, such as text-to-speech, a dictionary and multimedia glossary, and bookmarking. Some of these features could align with multiple guidelines and principles. According to CAST's suggestions for UDL implementation, text-to-speech could offer alternatives for visual information (checkpoint 1.3), support text decoding (checkpoint 2.3), and allow for construction and composition (checkpoint 5.2) depending on intended design needs. Thus, if researchers do not clarify the rationale for the design features and align them to specific checkpoints, it will be challenging for other researchers to replicate successful implementation components.

Patterns and Gaps Emerging from Aligned Checkpoints

Cross marks in Table 2 illustrate the alignment points between the intervention or instructional designs reported in the literature and UDL checkpoints. It is important to note again that for studies that did not explicitly reference checkpoints, we evaluated alignment between the described intervention or instructional designs and the description and suggestions for implementing each UDL checkpoint offered by CAST. Interested readers can refer to Supplemental Table 6 for the distribution patterns and numbers of checkpoints, guidelines, and principles between studies with explicit references to checkpoints and studies without explicit references but with assessed alignment to checkpoints. In this section, we reported on patterns and gaps that emerged from the synthesis of aligned checkpoints across all studies.

First, the number of design features aligned to checkpoints ranged from zero to 21, with most studies ($n = 22$) including designs aligned to four to 14 checkpoints that were distributed across all three principles. We were unable to code checkpoint alignment for five studies due to a lack of information on specific design features of interventions or instruction. The observed variations in the number of aligned checkpoints demonstrate the flexibility inherent in UDL implementation: interventions and instruction can draw on one, some, many, or all of the UDL checkpoints. On the other hand, such variations across UDL implementations seem to perpetuate the challenge of clearly defining and evaluating UDL as a framework for supporting all students. For example, there are no clear criteria for evaluating how well Zhang et al. (2021) and Marino et al. (2014) operationalized UDL, providing no guidance on how to interpret the vast difference in the number of checkpoints that were explicitly referenced in both studies.

Second, the largest number of interventions or instructional designs were aligned to checkpoints within the Representation principle ($n = 101$), followed by Action and Expression ($n = 68$) and Engagement ($n = 59$). Overall, the most aligned checkpoints were 1.3, 2.3, and 2.5, all corresponding to the Representation principle. The considerable number of aligned checkpoints within representation may reflect the original focus of UDL on designing accessible learning materials and formats through technology (Rose & Meyer, 2002) and a continuation of such focus over the history of UDL research. However, researchers (e.g., Hollingshead et al., 2022) have posited that research and practices on UDL have shifted from "accessibility" to "engagement" since CAST moved the Engagement principle to the forefront of the visualization of the framework (i.e., the UDL Guidelines Version 2.1; CAST, 2014). The lower frequency of checkpoints within Engagement indicates a lag in research that substantiates this shifting effort.

Third, collectively, previous studies investigated all checkpoints; however, most checkpoints ($n = 21$; 67.7%) were applied in less than ten studies, revealing the limited research investigating these checkpoints in the extant literature. For example, only one study included designs aligned to checkpoint 3.4 (maximize transfer and generalization). The uneven distribution of UDL checkpoints, guidelines, and principles has resulted in a lack of clarity regarding the role and evidence of lesser-applied elements of UDL. Even for multiple frequently applied checkpoints, the research base appears to be outdated. For example, any instructional design involving multimedia use aligns broadly with UDL checkpoint 2.5 (i.e., illustrated through multiple media), making it difficult to define and measure the contribution and effects of this checkpoint as part of UDL implementation.

RQ2: Presence of Theories and Extent of Theoretical Alignment in UDL Implementation

Presence of Theories in UDL Implementation at the Study Level

Only three studies explicitly referenced existing theories when describing the design of interventions or instruction (see Table 3). We distinguished how design features in these studies were linked to UDL and theories by identifying whether researchers used theories to justify the design process (i.e., direct theoretical guidance on designs), whether design features that were driven by theories happened to align with certain elements of UDL (i.e., indirect linkage between theories and designs aligned to UDL), or whether design features that were not driven by theories aligned to UDL checkpoints (i.e., no linkage between theories and designs aligned to UDL). The varying levels of linkage revealed different roles that theories played in UDL implementation.

Zhang et al. (2021) applied self-regulated learning (SRL; Schunk & Greene, 2018) theory to justify the design of instructional tools aligned to UDL. Specifically, the researchers explained how Zimmerman's (2000) SRL model guided the research team and a participating educator in co-creating a student self-assessment tool. The co-creation process and the tool development were driven by the educator's instructional needs that emerged from UDL implementation. However, the

Table 3 Presence of theories and alignment of the Universal Design for Learning (UDL) implementation to theories

Study	Theory Cited in the Study	Degree of Alignment in UDL Checkpoint Implementation to Theories									Measured Learning Outcome
		Representation			Action & Expression			Engagement			
		G1	G2	G3		G5	G6	G7	G8	G9	
		1.1 1.2 1.3	2.1 2.2 2.3 2.4 2.5	3.1 3.2 3.3 3.4	4.1 4.2	5.1 5.2 5.3	6.1 6.2 6.3 6.4	7.1 7.2 7.3	8.1 8.2 8.3 8.4	9.1 9.2 9.3	
Hall et al., 2015	Sociocultural Theory (Prior, 2006)	□	□ □	□	▣ □	▣	▣ ▣ ▣	▣	□	▣	Positive results: Increased reading scores
Kennedy et al., 2014	Cognitive theory of multimedia learning (Mayer, 2005)	▣		▣							Positive results: Increased vocabulary scores
Zhang et al., 2021	Self-regulated learning Theory (Zimmerman, 2000)									■	Null results: No increase in student engagement

Note. G = Guideline; The symbol □ indicates that the design feature associated with the specific UDL checkpoint was not linked to the theory referenced in the study. The symbol ▣ indicates an indirect link between the design feature aligned to the specific UDL checkpoint and the theory referenced in the study. The symbol ■ indicates that the design feature associated the specific UDL checkpoint was guided by the theory referenced in the study.

theory application in this study was limited to one design aligned to checkpoint 9.3 (as illustrated through the solid square in Table 3) and one element of Zimmerman's SRL model.

Two studies integrated existing theories, UDL, and other strategies into the design of the investigated instructional practice or tool. Kennedy et al. (2014) designed Content Acquisition Podcasts (CAPs), a multimedia-based instructional tool, in accordance with UDL principles and Mayer (2005). According to the researchers, Mayer's CTML principles guided all production decisions for designing CAPs; UDL was also considered as CAPs provided an alternative mode of presenting instruction, such as using visuals (aligned to checkpoints 1.2 and 2.5; as illustrated through gray squares in Table 3), and a flexible tool for students to engage in learning (broadly linked to the Engagement principle but not specific to any checkpoint). In this case, the design features of CAPs were directly justified by CTML rather than UDL; however, two features corresponded to UDL checkpoints, thus showing an indirect link to UDL.

Similarly, Hall et al. (2015) integrated sociocultural theory, UDL, curriculum-based measurement (CBM), and reciprocal teaching strategies into a digital reading environment. These four elements converged to build a conceptual foundation for the environment embedded with 13 design features aligned to UDL checkpoints. Most features related to the Representation principle, such as text-to-speech, were not directly linked to the sociocultural theory of literacy cited in the study (as illustrated through blank squares in Table 3). Researchers used the theory to justify the design of an online forum to facilitate community building and communication among students and teachers. Some of these designs aligned with the UDL checkpoints that provide suggestions on fostering community and communication (e.g., 5.1, 8.3). As such, these designs were indirectly linked to the theory rather than being directly informed or justified by the theory.

The identified linkages among intervention or instructional designs, UDL, and theories in the limited existing research point to a significant gap in applying

theories to guide the UDL design process. Of the 28 studies without explicit references to theories, only a few studies applied UDL in a systematic way. For example, Root et al. (2020) applied UDL to customize existing components of a mathematics evidence-based practice (EBP) to tailor the EBP to the needs of individual SWDs. This approach was in line with the systematic implementation of UDL proposed by Cook and Rao (2018) in which researchers or practitioners use UDL guidelines and checkpoints to adapt an existing EBP while maintaining its core components. Such an approach provides a methodologically sound structure to systematically implement and measure the effect of UDL when applied to effective practices provided that researchers can distinguish the effects of the practice with and without UDL-based adaptations (Cook & Rao, 2018). Nevertheless, most remaining studies did not show a clear structure for using the framework due to the absence of theoretical guidance on design processes or UDL-guided adaptations to effective practices. The lack of structures portrays the challenge of systematically implementing and measuring UDL as a design framework.

Theoretical Alignment of Checkpoint Implementation

Due to the limited presence of theories in individual studies, we further analyzed whether the implementation of checkpoints across studies, collectively, aligned with existing theories. Our intent was to identify emergent themes or practices across studies that might echo fundamental tenets of learning or instructional design theories, thus offering new insights into UDL implementation driven by theories. We focused on checkpoints 1.3, 6.2, and 7.1 as examples given that they represent the most-applied checkpoint within each principle (see Supplemental Table 7 for detailed descriptions of design features aligned to these checkpoints). Due to space limitations, we grounded the analysis for each checkpoint within a single established theory (see Supplemental Table 1 for all theories referenced on the CAST website).

Aligning with checkpoint 1.3 (i.e., offer alternatives for visual information), the text-to-speech function emerged as the most-applied design feature, which was frequently embedded in digital tools or platforms, in the existing UDL literature. Other design features included read-aloud support, animations, illustrations complementing text, pictorial and verbal explanations, narrations with corresponding visuals, alternative text, and pre-recorded lectures. Most of these designs were essential for ensuring the accessibility of visual information, especially for students with visual disabilities. Delving further into theories, a frequently referenced theory that supports the development of this checkpoint is Mayer (2005). In essence, CTML was developed based on the assumption that humans process information through dual channels for verbal and visual information, each channel has limited processing capacity at one time, and meaningful learning involves active cognitive processing in both channels (Mayer, 2005). CTML and numerous validated design principles derived from its underlying assumptions offer extensive guidance on designing instruction to facilitate intricate information processing (e.g., selecting, organizing, and integrating information in dual channels) and knowledge construction (e.g., Sweller et al., 2019; Mayer & Moreno, 2003). For most empirical designs aligned to checkpoints 1.3 and 1.2 (i.e., offer alternatives for auditory information), the

researchers acknowledged the importance of offering both visual and verbal information for accessibility but fell short of addressing the complexities of more intricate instructional designs.

Design features aligned to checkpoint 6.2 (i.e., support planning and strategy development) varied across studies. Commonly applied designs included providing prompts, hints, or guidance for students to apply learning strategies. Such guidance included modeling strategy use, offering graduated scaffolding, and implementing think-aloud strategies through varied resources and tools (e.g., pedagogical agents, teachers, and instructional videos). Such arrayed designs illustrate the broadness of this checkpoint as it integrates various pedagogical approaches and theories that substantiate its effectiveness. The most frequently referenced approach is Self-Regulated Strategy Development (SRSD; Harris & Graham, 1999), along with one of its theoretical groundings—SRL theories (Pintrich, 2000). In a nutshell, SRSD involves six stages that guide students' acquisition and application of a strategy (i.e., develop background knowledge, discuss it, model it, memorize it, support it, and independent performance; Harris & Graham, 1999). Although there are different models, SRL outlines sequential phases of a learning process, such as task planning, monitoring, strategy use, reaction, and reflection (Greene, 2017; Pintrich, 2000). From both SRSD and SRL perspectives, students' strategy development interacts with other stages or phases of a learning process; thus, this begs the question as to whether implementing checkpoint 6.2 requires considerations for checkpoints or elements associated with other stages or phases of SRSD or SRL.

In terms of theory, the development of checkpoint 7.1 (optimize individual choice and autonomy) was largely reliant upon SDT. As a macro theory focused on human motivation and personality, SDT explains three basic psychological needs for autonomy, competence, and relatedness (Deci & Ryan, 2002). Checkpoint 7.1 highlights providing options and choices to promote autonomy. In the studies we reviewed, such options and choices included providing students opportunities to learn at their own pace, design learning tasks, exercise autonomy over task sequences, make choices in learning topics, have the freedom to choose tasks with varying levels of challenge, select types of rewards or recognition, or use different tools for production). However, checkpoint 7.1 overlaps with other UDL checkpoints and guidelines to a great extent. Each UDL guideline starts with "provide options" in its descriptor, and corresponding checkpoints further demonstrate several ways to provide options and choices to meet diverse student needs. For example, in Dalton et al. (2011), students were provided with autonomy over whether to type or audio-record responses and choices to choose visual text display and text-to-speech output. These design features align with checkpoint 7.1 due to identified student choices and autonomy. Additionally, they can be aligned to other checkpoints given specific aspects of those choices (e.g., checkpoint 4.1 varies in the methods for response and navigation). Thus, the implementation of checkpoint 7.1 illustrates a dilemma in which disconnections may exist among checkpoints grounded in different elements of the same theory while overlaps may emerge when implementing checkpoints that share similar theoretical groundings.

Discussion

As UDL has entered its fourth decade of development and research, there are ongoing critiques of UDL for the lack of clarity in its definitions and implementation (Edyburn, 2021; Matthews et al., 2023; Smith et al., 2019). The field of UDL research seems to stagnate in ongoing debates without concrete solutions (Hollingshead et al., 2022). To shed some light on future UDL research, we analyzed how researchers implemented UDL checkpoints, guidelines, and principles in the extant literature with a focus on unraveling the challenges with UDL implementation. Additionally, we assessed the extent to which the implementation process was guided by existing theories within and across individual studies with the intent to provide insights into systematic UDL implementation aligned to theories. The challenges emerging from our analyses that stymie UDL research encompass several interrelated aspects: the absence of explicit alignment between UDL checkpoints and the design features of interventions or instruction investigated in the literature, an uneven distribution of implemented checkpoints and corresponding guidelines, confusion derived from the overlap among multiple checkpoints and guidelines, and the lack of theoretical guidance on UDL design and implementation. Due to the limited space, we offer the following two recommendations for strengthening methodological and theoretical aspects of UDL implementation as starting points.

Recommendation One: Strengthen Research Base and Specify UDL Checkpoints

Our review revealed that a limited number of studies ($n = 8$) explicitly referenced UDL checkpoints in the design of interventions or instruction. This lack of clear alignment illustrated confusion regarding what constitutes UDL implementation. In particular, numerous instructional designs can be loosely linked to UDL due to the broadness of the language describing UDL checkpoints, guidelines, and principles (e.g., checkpoint 2.5 Illustrate through multiple media). Even in studies with explicit references to checkpoints, confusion emerged because some instructional designs corresponded to different guidelines and principles. When one design element can be aligned to multiple checkpoints, it is unclear whether it was intentionally designed to do so (e.g., Root et al., 2020; King-Sears & Johnson, 2020) or if it was an artifact of the overlap among multiple UDL checkpoints. Either of these cases complicates the way researchers define and measure UDL implementation.

In the first case, researchers may argue that intentionally designing components of an intervention or instruction that align with multiple UDL checkpoints showcases the flexibility of UDL implementation. However, such intentional designs add another layer of complexity to the inconsistent approaches to UDL implementation, namely varied combinations of checkpoints and guidelines applied across studies, as shown in both current and prior reviews (e.g., Ok et al., 2017). The complexity has resulted in a stagnation of research on UDL effectiveness, even though scholars (e.g., Edyburn, 2010) have long called for the need to investigate and measure

essential, replicable elements of systemic UDL implementation. The challenge of identifying essential elements was further complicated by the uneven distribution of checkpoints implemented across studies and the fact that the literature base for multiple UDL checkpoints is outdated or disconnected from the framework (Matthews et al., 2023). Thus, more research is needed to investigate these checkpoints to determine their contribution to UDL implementation.

Regarding RQ2, results from our analysis on whether the implementation of checkpoints across individual studies exhibited alignment with existing theories revealed a potential source of overlap among checkpoints, which subsequently resulted in confusion surrounding UDL implementation. We found that the most-applied checkpoints within each principle, alongside multiple other checkpoints, were developed based on certain elements of existing theories (e.g., CTML, SRL). These checkpoints are naturally connected to other checkpoints developed based on different elements of the same theory. For example, the development of checkpoints 6.4 and 9.3 was grounded in SRL, and checkpoints 1.2, 1.3, and 2.5 were rooted in CTML. Both sets of checkpoints exhibit theoretical connections despite that individual checkpoints correspond to different guidelines and/or principles. Existing theories suggest relationships among design features of intervention or instruction, which can guide the systematical implementation of UDL that highlights connections among checkpoints, guidelines, and principles. These considerations are absent from the existing UDL research, thus requiring further investigation.

These intricate challenges with UDL implementation point to a pressing need for more foundational research on UDL checkpoints and guidelines that rely upon outdated research or were less frequently applied in previous UDL research to establish a stronger research base for UDL implementation (Beerwart, 2018; Matthews et al., 2023). For example, there is a need to update the research base for such checkpoints as 2.5 and 4.2 (which only have "face validity" indicated on the CAST website). This foundational research will need collective efforts from the UDL community to thoroughly examine and synthesize up-to-date literature on checkpoints and guidelines. Second, reframing UDL checkpoints and guidelines by clarifying their relationships based on their theoretical groundings seems necessary for reducing confusion derived from the overlap among checkpoints and guidelines. These efforts could potentially enhance the coherence and logical structure of the UDL framework. Further, updating and reframing checkpoints allow researchers to implement checkpoints in a coherent way that aligns with underlying theories and measures the effects of specified checkpoints more systematically, thus testing the efficacy of each checkpoint through replicable, accumulating evidence.

With an updated framework in place, it is imperative for researchers to specify checkpoints and corresponding guidelines as a more transparent approach to operationalizing UDL, which aligns with the suggestion provided by other UDL researchers (e.g., Rao et al., 2020). More importantly, researchers should provide rationales or theoretical explanations for the designs that are essential to their UDL operationalization. This requires more efforts to substantiate the effectiveness of these checkpoints as indispensable components of UDL implementation rather than claiming that UDL is scientifically validated given the extensive research supporting its checkpoints, which were oftentimes substantiated outside the context of UDL implementation (Beerwart, 2018; Matthews et al., 2023). These recommendations

are in line with ongoing calls for addressing the paucity of research investigating the effects of UDL as a single construct, namely a coherent design framework (e.g., Basham et al., 2020; Edyburn, 2021), as well as provide insights into systemically implementing UDL guided by theories.

Recommendation Two: Establish Systematic UDL Implementation Guided by Theories

Although researchers need to conduct more foundational research to strengthen the theoretical and empirical base of UDL, the framework has its merits in providing a way of thinking that resists deficit views of SWDs and supports learner variability from the onset of instruction designs (Waitoller & Thorius, 2016). It is important to acknowledge that implementing and evaluating the effects of UDL as a complex design framework is by no means straightforward. Researchers in the UDL community proposed approaches to systematically implementing UDL to address the complexity, such as Cook and Rao's (2018) suggestion on utilizing the framework to adapt elements of an EBP to meet individual learners' needs. However, most studies included in this review did not employ this systemic approach to investigating UDL adaptations; instead, UDL was mainly used to design a tool, strategy, or curriculum with significant variations in implemented checkpoints and guidelines.

From the instructional design perspective, UDL scholars have proposed that systemic UDL implementation should encompass both *proactive* and *iterative* instructional design processes (e.g., Basham et al., 2020; Smith et al., 2019). Proactive design processes require researchers or educators to anticipate that learners will vary in why, what, and how they learn (Meyer et al., 2014); most existing research efforts have emphasized such processes. Iterative design processes involve ongoing decision-making to address the emergent learning needs of individual learners when they interact with a specific learning context (Basham et al., 2020). However, research investigating iterative design processes as a part of UDL implementation is largely absent from the current literature. To address this gap, we suggest that researchers attend to theoretical guidance on iterative design processes given the utility of established theories for explaining dynamic interactions among elements of a complex system and causal mechanisms existing in the system that may lead to desired outcomes (Braithwaite et al., 2018; Nilsen, 2015).

We acknowledge that the comprehensive nature of the UDL framework encompasses multiple strategies, methodologies, and theoretical underpinnings, limiting the development and application of a cohesive theory for UDL implementation. Thus, it might be impracticable to apply theories to justify a large combination of varied design features when implementing UDL. In this regard, we suggest that researchers provide theoretical explanations for major design decisions, align designs to checkpoints and guidelines, and explicate potential interacting influences from UDL-aligned instructional designs. Figure 2 illustrates a process of designing specific components of an intervention or instruction guided by validated theories as a part of UDL implementation. Given the design process is situated within a given context, we adopted Dewey's (1933) systematic approach to inquiry in social research that involves a

five-step process of reflective decision-making, which includes recognizing problems, considering the nature of the problem, suggesting solutions, considering effects of solution, and taking action. According to Dewey, research inquiry should be situated within a given context where researchers engage in a cyclical process of understanding unknowable problems, designing and testing solutions, and generating new understandings that result from taking action (Morgan, 2014).

Guided by Dewey's process of inquiry, our proposed design approach first suggests that UDL implementation be situated within a specific context with unique historical and cultural factors (as depicted at the top of Fig. 2). Attention to the context of an implementation effort opens space for active participation from stakeholders (e.g., educators, administrators, researchers), which is an essential factor that facilitates sustainable implementation (Braithwaite et al., 2018). Moreover, active stakeholder engagement ensures that the implementation is informed by data collected via various sources within the context and tailored to contextual characteristics, thus maintaining flexibility in how stakeholders apply UDL. Specifically, the need to factor in context begins with identifying problems (e.g., barriers to student learning) and implementation needs (e.g., supporting educators in designing instruction to address the barriers). Then, information related to contextual problems helps identify facilitators or impediments that may impact the design of solutions to address identified problems (e.g., [mis]alignment to existing practices, [in]sufficient system support; May et al., 2016).

This process of designing solutions involves iterations of design, application, and evaluation of interventions or instructional practices that should be guided by theories and driven by target student learning outcomes (depicted at the bottom of Fig.2). After determining specific student learning outcomes, researchers and instructional designers should attend to validated theories to guide the design of interventions or instructional practices. Researchers and instructional designers should justify why the major or core designs of an intervention or instruction can lead to increased student learning outcomes. Further, they can consider whether and how the core designs guided by theories align with UDL guidelines and checkpoints. As many educational theories were developed based on insufficient considerations for the experience of SWDs (Greene, 2022; Emery & Anderman, 2020), it is imperative for researchers to consider how these theoretically guided designs can be enhanced by considering UDL (e.g., enhancing accessibility in core designs). To improve the replicability of the design features, researchers should clearly articulate which and how UDL checkpoints and guidelines are applied to the theory-driven designs. To illustrate, Zhang et al. (2021) could have designed instructional practices integrating core elements of an SRL model to enhance student engagement. Instead, researchers only focused on one design that reflected the last phase of Zimmerman's SRL model by implementing checkpoint 9.3, thus missing the opportunity to explore the interrelationship between checkpoint 9.3 and other checkpoints grounded in SRL theories (e.g., 6.1, 6.2, 9.1, and 9.2).

In line with the systematic application of UDL to EBPs proposed by Cook and Rao (2018), the iterative process prioritizes designing instructional practices guided by existing theories rather than relying upon the framework to select a set of discrete designs. Thus, this process indicates a priori theoretical considerations for UDL-aligned designs, through which stakeholders implement UDL checkpoints or
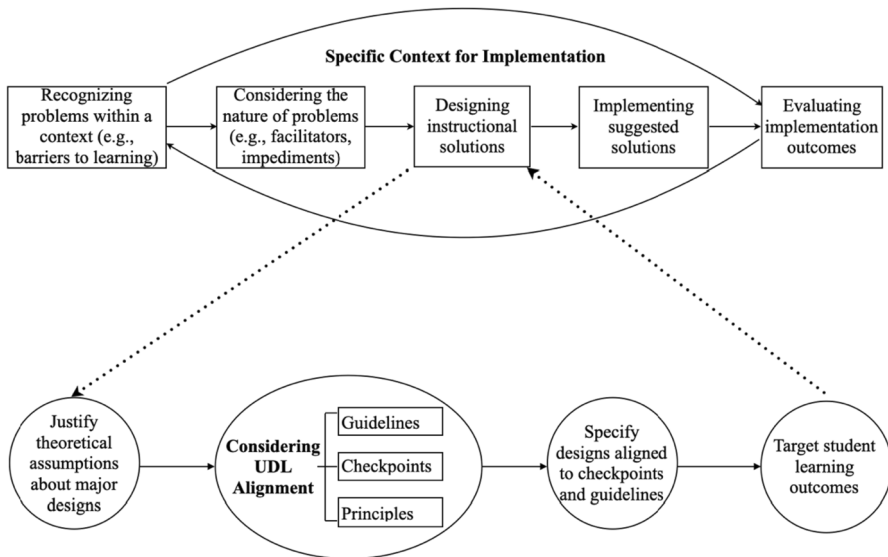
Fig. 2 Theoretical approaches to designing intervention or instruction aligned to the Universal Design For Learning (UDL). This figure illustrates an iterative process of implementing Universal Design for Learning (UDL). The process at the upper level was adopted based on Dewey (1933) systematic approach to inquiry as a five-step reflective process of decision-making in social research. This process highlights active participation from stakeholders (depicted at the top level) to ensure implementation tailored to contextualized needs. The process further incorporates an iterative process of designing, implementing, and evaluating interventions or instructional practices guided by theories (depicted at the bottom level)

guidelines as interdependent components, defined by their theoretically informed interrelations in contributing to successful design efforts. When these interrelations emerge in the form of recurring patterns, they will help clarify the role of the UDL framework in iteratively designing interventions or instruction. Moreover, the iterative process of implementing and evaluating designed solutions provides data necessary for informing ongoing improvement in implementation by adapting interventions or instruction to meet emergent instructional or learning needs. Overall, this iterative process of designing, implementing, and evaluating solutions as part of the large implementation effort prioritizes the application of UDL checkpoints tailored to the specific context while providing a structure for measuring the UDL-based adaptions to theoretically guided designs. In this article, we focused on unpacking "designing instructional solutions" as a core step of an iterative UDL implementation process. More research is needed to unpack other steps, establish guidance, and develop tools to facilitate the complex implementation process with coherency and clarity.

Limitations

There are two limitations to this review that warrant interpreting the results with caution. First, we focused on peer-reviewed studies that used research methods

with quantitative components to evaluate student learning outcomes in preK-12 settings. Although we selected these inclusion criteria to examine UDL implementation as reflected in measured learning outcomes, gray literature (e.g., dissertations) or studies with qualitative research designs might provide different, useful information on UDL implementation. In the future, researchers should consider including gray literature, qualitative studies, and studies focused on other settings (e.g., higher education) to provide a more comprehensive picture of UDL implementation.

A second limitation relates to the analysis of the alignment of intervention or instructional designs to UDL checkpoints reported in studies where researchers did not specify checkpoints/guidelines when describing their designs. It is possible the authors of studies without explicit references to checkpoints had different intentions and would have coded them differently, due to the broadness of language describing UDL checkpoints, guidelines, and principles. Despite substantial agreement between the two coders, the results of alignment for studies without explicit references should be interpreted with caution. This limitation speaks directly to the importance of specifying UDL checkpoints and corresponding guidelines in future research. Therefore, it was essential to align the designs reported in all reviewed studies to UDL checkpoints through our coding procedure. This provided a comprehensive understanding of UDL-guided designs of interventions or instructions with varying levels of specificity regarding how researchers referenced checkpoints. Further, we have transparently described our data collection and analysis so that other interested researchers in the field can replicate this study.

Conclusion

UDL researchers have emphasized the importance of acknowledging challenges with working in transdisciplinary disciplines while appreciating opportunities that UDL affords for stimulating interconnected work (e.g., Rappolt-Schlichtmann et al., 2012). Thus, as UDL enters its 4th decade of research and development, it is imperative that researchers in the field acknowledge and address ongoing critiques of the framework moving forward. The findings of this literature review unveiled the absence of explicit alignment of instruction or intervention in previous research to UDL checkpoints and the lack of theoretical guidance on the design and implementation processes as prominent factors that may have caused these critiques. We suggested specifying checkpoints in UDL implementation, strengthening the research base for checkpoints, and establishing a theoretically guided design process that is embedded within an iterative approach to systematically implementing UDL. Moreover, we acknowledged that more research is needed to investigate all steps of the proposed iterative approach and examine the efficacy of UDL as a coherent framework for promoting inclusive learning experiences for all.

References

(Note: References of all included studies can be found in Supplemental Table 3)

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*, 171–185. https://doi.org/10.1007/s10984-011-9090-2

Al-Azawei, A., Serenelli, F., & Lundqvist, K. (2016). Universal design for learning (UDL): A content analysis of peer reviewed journals from 2012 to 2015. *Journal of the Scholarship of Teaching and Learning, 16*(3), 39–56. https://doi.org/10.14434/josotl.v16i3.19295

Basham, J. D., Gardner, J. E., & Smith, S. J. (2020). Measuring the implementation of UDL in classrooms and schools: Initial field test results. *Remedial and Special Education, 41*(4), 231–243. https://doi.org/10.1177/0741932520908015

Beerwart, A. A. (2018). *An analytical integrative review of universal design for learning's effectiveness for K-12 students (publication no. 10935531)* Doctoral dissertation, Regent University, ProQuest Dissertations & Theses Global.

Bernacki, M. L. (2018). Examining the cyclical, loosely sequenced, and contingent features of self-regulated learning: trace data and their analysis. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of Self-Regulated Learning and Performance.* Routledge.

Boysen, G. A. (2021). Lessons (not) learned: The troubling similarities between learning styles and universal design for learning. *Scholarship of Teaching and Learning in Psychology*. https://doi.org/10.1037/stl0000280

Braithwaite, J., Churruca, K., Long, J. C., et al. (2018). When complexity science meets implementation science: A theoretical and empirical analysis of systems change. *BMC Medicine, 16*, 63. https://doi.org/10.1186/s12916-018-1057-z

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). *Universal design for learning guidelines version 2.1*. CAST.org.

CAST. (2018a). *Universal design for learning guidelines version 2.2*. CAST.org.

CAST. (2018b). *UDL and the learning brain*. CAST.org https://www.cast.org/products-services/resources/2018/udl-learning-brain-neuroscience

Cook, S. C., & Rao, K. (2018). Systematically applying UDL to effective practices for students with learning disabilities. *Learning Disability Quarterly, 41*, 179–191. https://doi.org/10.1177/0731948717749936

Courey, S. J., Tappe, P., Siker, J., & LePage, P. (2013). Improved lesson planning with universal design for learning (UDL). *Teacher Education and Special Education, 36*(1), 7–27. https://doi.org/10.1177/0888406412446178

Deci, E. L., & Ryan, R. M. (2002). *Handbook of self-determination research*. University of Rochester Press.

Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process*. Heath & Co Publishers.

Edyburn, D. (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*, 33–41. https://doi.org/10.1177/07319487100330010

Edyburn, D. L. (2021). Ten years later: Would you recognize universal design for learning if you saw it? *Intervention in School and Clinic, 56*(5), 308–309. https://doi.org/10.1177/1053451220963114

Emery, A., & Anderman, L. H. (2020). Using interpretive phenomenological analysis to advance theory and research in educational psychology. *Educational Psychologist, 55*(4), 220–231. https://doi.org/10.1080/00461520.2020.1787170

Every Student Succeeds Act of 2015, (2015) 20 U.S.C. § 4104 et seq.

Fornauf, B. S., & Erickson, J. D. (2020). Toward an inclusive pedagogy through universal design for learning in higher education: A review of the literature. *Journal of Postsecondary Education and Disability, 33*(2), 183–199.

Greene, J. A. (2017). *Self-regulation in education*. Routledge.

Greene, J. A. (2022). What can educational psychology learn from, and contribute to, theory development scholarship? *Educational Psychology Review, 34*(4), 3011–3035. https://doi.org/10.1007/s10648-022-09682-5

Harris, K. R., & Graham, S. (1999). Programmatic intervention research: Illustrations from the evolution of self-regulated strategy development. *Learning Disability Quarterly, 22*(4), 251–262. https://doi.org/10.2307/1511259

Higher Education Opportunity Act of 2008, (2008) 20 U.S.C. 1001 et seq.

Hollingshead, A., Lowrey, K. A., & Howery, K. (2022). Universal design for learning: When policy changes before evidence. *Educational Policy, 36*, 1135–1161. https://doi.org/10.1177/0895904820951120

Irbe, A. G. (2016). *Application of universal design for learning in corporate technical training design: A quantitative study (publication no. 10014050)*, Doctoral Dissertation, Capella University. ProQuest Dissertations & Theses Global.

Kennedy, J., Missiuna, C., Pollock, N., Wu, S., Yost, J., & Campbell, W. (2018). A scoping review to explore how universal design for learning is described and implemented by rehabilitation health professionals in school settings. *Child: Care, Health and Development, 44*(5), 670–688. https://doi.org/10.1111/cch.12576

Kennedy, M. J., Thomas, C. N., Meyer, J. P., Alves, K. D., & Lloyd, J. W. (2014). Using evidence-based multimedia to improve vocabulary performance of adolescents with LD: A UDL approach. *Learning Disability Quarterly, 37*(2), 71–86. https://doi.org/10.1177/0731948713507262

King-Sears, M. (2009). Universal design for learning: Technology and pedagogy. *Learning Disability Quarterly, 32*(4), 199–201. https://doi.org/10.2307/277403

King-Sears, M. E., Stefanidis, A., Evmenova, A. S., Rao, K., Mergen, R. L., Owen, L. S., & Strimel, M. M. (2023). Achievement of learners receiving UDL instruction: A meta-analysis. *Teaching and Teacher Education, 122*, 103956. https://doi.org/10.1016/j.tate.2022.103956

Kortering, L. J., McClannon, 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/10.1177/074193250731402

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159–174. https://doi.org/10.2307/2529310

Mace, R. (1997). *What is universal design.* The Center for Universal Design at NCSU. https://disabilityandmultimodality.wordpress.ncsu.edu/universal-design-ud/

Marino, M. T., Black, A. C., Hayes, M. T., & Beecher, C. C. (2010). An analysis of factors that affect struggling readers' achievement during a technology-enhanced STEM astronomy curriculum. *Journal of Special Education Technology, 25*(3), 35–47. https://doi.org/10.1177/0162643410025003

Matthews, S., Cavanaugh, C., & Wilson, P. H. (2023). Multiple means of representation? A critical analysis of universal design for learning checkpoint 1.2. *Policy Futures in Education, 21*(2), 190–204. https://doi.org/10.1177/14782103221097514

May, C. R., Johnson, M., & Finch, T. (2016). Implementation, context and complexity. *Implementation Science, 11*, 141. https://doi.org/10.1186/s13012-016-0506-3

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48). Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.004

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 43–52. https://doi.org/10.1207/S15326985EP3801_6

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemical Medicine, 22*, 276–282.

Meyer, A., Rose, D. H., & Gordon, D. (2014). *Universal design for learning: Theory and practice*. CAST Professional Publishing.

Miller, D. M., Scott, C. E., & McTigue, E. M. (2018). Writing in the secondary-level disciplines: A systematic review of context, cognition, and content. *Educational Psychology Review, 30*, 83–120. https://doi.org/10.1007/s10648-016-9393-z

Morgan, D. L. (2014). Pragmatism as a paradigm for social research. *Qualitative Inquiry, 20*(8), 1045–1053. https://doi.org/10.1177/1077800413513733

Nilsen, P. (2015). Making sense of implementation theories, models and frameworks. *Implementation Science, 10*(1), 53. https://doi.org/10.1186/s13012-015-0242-0

Ok, M. W., Rao, K., Bryant, B. R., & McDougall, D. (2017). Universal design for learning in pre-k to grade 12 classrooms: A systematic review of research. *Exceptionality, 25*(2), 116–138. https://doi.org/10.1080/09362835.2016.1196450

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—a web and mobile app for systematic reviews. *Systematic Reviews, 5*, 1–10. https://doi.org/10.1186/s13643-016-0384-4

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *International Journal of Surgery, 88*, 105906. https://doi.org/10.1016/j.ijsu.2021.105906

Paul, J., Khatri, P., & Kaur Duggal, H. (2023). Frameworks for developing impactful systematic literature reviews and theory building: What, why and how? *Journal of Decision Systems*, 1–14. https://doi.org/10.1080/12460125.2023.2197700

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation*, pp. 452–502

Prior, P. (2006). A sociocultural theory of writing. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 54–66). The Guilford Press.

Rao, K., Ok, M. W., & Bryant, B. R. (2014). A review of research on universal design educational models. *Remedial and Special Education, 35*, 153–166. https://doi.org/10.1177/0741932513518980

Rao, K., Ok, M. W., Smith, S. J., Evmenova, A. S., & Edyburn, D. (2020). Validation of the UDL reporting criteria with extant UDL research. *Remedial and Special Education, 41*(4), 219–230. https://doi.org/10.1177/07419325198477

Rappolt-Schlichtmann, G., Daley, S. G., & Rose, L. T. (2012). *A research reader in universal design for learning*. Harvard Education Press.

Rose, D. H., & Meyer, A. (2002). *Teaching every student in the digital age: Universal design for learning*. Association for Supervision and Curriculum Development.

Saldaña, J. (2020). Qualitative data analysis strategies. In Patricia Leavy (Ed.), *The Oxford handbook of qualitative research* (2nd ed., pp. 876–911). Oxford University Press.

Schunk, D. H., & Greene, J. A. (Eds.). (2018). *Handbook of Self-Regulation of Learning and Performance* (2nd ed.). Routledge.

Smith, S. J., Rao, K., Lowrey, K. A., Gardner, J. E., Moore, E., Coy, K., Marino, M., & Wojcik, B. (2019). Recommendations for a national research agenda in UDL: Outcomes from the UDL-IRN preconference on research. *Journal of Disability Policy Studies, 30*(3), 174–185. https://doi.org/10.1177/1044207319826219

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*, 261–292. https://doi.org/10.1007/s10648-019-09465-5

Waitoller, F. R., & Thorius, K. A. K. (2016). Cross-pollinating culturally sustaining pedagogy and universal design for learning: Toward an inclusive pedagogy that accounts for dis/ability. *Harvard Educational Review, 86*(3), 366–389. https://doi.org/10.17763/1943-5045-86.3.366

Webb, K. K., & Hoover, J. K. (2015). Universal design for learning (UDL) in the academic library: A methodology for mapping multiple means of representation in library tutorials. *College & Research Libraries, 76*(4), 537–553 http://hdl.handle.net/10342/4970

Zhang, L., Jackson, H. A., Yang, S., Basham, J. D., Hunt, C. L., & Carter, R. A., Jr. (2021). Codesign learning environments guided by the framework of UDL: A case study. *Learning Environments Research, 25*, 379–397. https://doi.org/10.1007/s10984-021-09364-z

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press.

Authors and Affiliations

Ling Zhang[1] ⬛ · **Richard Allen Carter Jr[2]** ⬛ · **Jeffrey A. Greene[1]** ⬛ ·
Matthew L. Bernacki[1] ⬛

✉ Ling Zhang
ling.zhang@unc.edu

[1] School of Education, University of North Carolina at Chapel Hill, 100 E Cameron Ave, Chapel Hill, NC, USA

[2] College of Education, University of Wyoming, Laramie, WY, USA