Hershock, C., Stimson, J., Pottmeyer, L. O., Melville, M. C., Harrell, J., Weiss, E. D., Rodriguez, M., le Blanc, S., & Adams, A. (2022). Thin-Sliced Embedded Assessment (T-SEDA): Measuring impacts of educational development workshops on participants' learning gains. *Journal on Centers for Teaching and Learning, 14*, 58-89.

Thin-Sliced Embedded Direct Assessment (T-SEDA): Measuring Impacts of Development Workshops on Participants' Learning Gains

Chad Hershock, Jacqueline Stimson, Laura Ochs Pottmeyer, Michael C. Melville, Jessica Harrell, Emily Daniels Weiss, Marisella Rodriguez, Sophie le Blanc, & Alexis Adams

How can educational developers best formatively assess impacts of their services? Standard practices tend to rely on indirect measures, such as counts of participants and feedback surveys. This paper responds to recent calls for more robust approaches directly measuring outcomes. We describe how to implement a new, transferable, evidence-based approach for directly measuring instructors' learning gains within and across educational development workshops: the Thin-Slice Embedded Direct Assessment (T-SEDA) Process. Although this approach does not measure long-term retention of learning or effects on future teaching behaviors, it significantly enhanced our toolkit for formatively assessing educational development workshops. Through case studies, we illustrate principles underlying our approach, impacts on instructors' learning, and how we iteratively refine our programs and practices using the T-SEDA process. We also discuss lessons learned for fostering an inclusive, collaborative culture of formative assessment among educational developers.

Introduction

In Educational Development (EdDev), we value and promote data-informed practices (Lovett & Hershock, 2020; POD 2018) to improve teaching and learning as well as EdDev programs and services. Because group programming is an essential component of this work (Gillespie & Robertson, 2010), we sought a way to more systematically and directly assess *participants' learning gains regarding pedagogical knowledge and skills*, specifically within and among short, independent, single-session workshops. We see these gains as a first step to better understanding how EdDev workshops

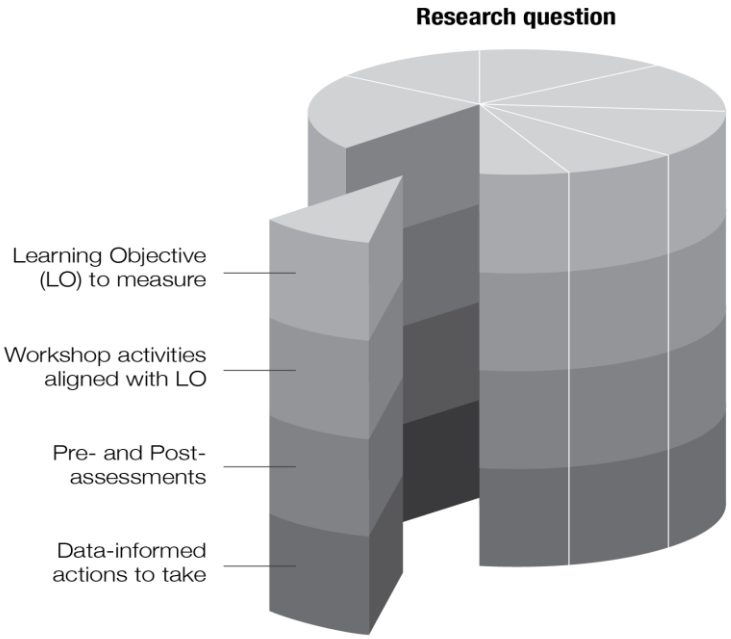
might impact instructor behaviors and/or student learning. Specifically, the context within which we implement single-session group programs poses significant challenges for any assessment plan attempting to directly measure participants' learning. We developed and piloted a new assessment paradigm to mitigate these challenges and provide actionable data on participants' learning within *and* across workshops.

Consistent with standard practice at most Centers for Teaching and Learning (CTLs), we have most typically evaluated our workshops via attendance and feedback surveys (Beach et al., 2016; Haras et al., 2017). While certainly helpful, survey data relying on self-reports do not necessarily provide valid and reliable formative assessments of outcomes (e.g., Ebert-May et al., 2011). Because attendance and survey data alone are insufficient to inform our future practice, we wanted to develop a better way to measure the extent to which attendees are achieving our learning objectives within and across these programs. Additionally, recent calls (see references above; Meizlish et al., 2018; POD 2018;) prompted us to find new ways to assess our CTL's impact. Our CTL director also challenged us to find a way to formatively yet directly measure our workshops' impacts in a sustainable, actionable way.

In 2018, we piloted a new strategy designed to effectively measure achievement of learning objectives (LOs) and to address the challenges associated with formative assessment *within and across workshops*. These challenges include participants entering each session with different prior knowledge and motivation as well as our desire to maintain a culture of learning and formative development (rather than summative evaluation of participants' or facilitators' skills). Time constraints also uniquely impact what is practical and possible to measure. For example, the instrument needed to assess 3-4 objectives in a 90-min workshop would take too much time, therefore reducing the time for actual workshop content. At the same time, delayed, online post assessments have inadequate response rates. We designed our new assessment approach to address these challenges while still obtaining quality data to inform iterative refinement of individual workshops. Additionally, *across workshops*, we want to identify which design strategies are most effective and generalizable and what patterns exist in attendees' learning gains. For example, which LOs are consistently met or not in workshops? For what are workshops best suited in terms of LOs? Answering these questions would not only enhance workshop design across our CTL's programming, but also inform strategic planning regarding what types of group programs to offer to best meet particular EdDev needs.

We call the assessment approach we designed the "Thin-Slice Embedded Direct Assessment Process," or the T-SEDA Process (Figure 1). The T-SEDA Process begins with collecting learning gains data within individual workshops. We do this by targeting no more than two LOs per workshop event.

Figure 1
The Relationship Between a Thin-Slice and the T-SEDA Process



Note. A thin-slice from a single workshop event includes the: (1) 1-2 learning objectives targeted by the assessment; (2) embedded pre-assessment items; (3) workshop content aligned with the targeted learning objectives; (4) embedded post-assessment items; and (5) actions resulting from the data (Figure 2). Collecting thin-slices across time and workshop events allows educational developers to explore larger-grained questions. For example, do 90-minute workshops enhance both instructors' lower- and higher-order cognitive skills associated with teaching? Here, we illustrate the aggregation of thin-slices targeting LOs focused on comprehension (lower-order skills) and analysis/application (higher-order skills).

After identifying LOs, we design and implement short, pre/post assessment instruments during the workshop to directly measure knowledge and/or skills aligned with the targeted LOs. These assessment instruments can be completed in five minutes and contain 1-4 items each, often including authentic tasks (e.g., suggest how to improve a lecture slide for learning). Hereafter, *thin-slice* refers to these targeted pre/post assessment instruments, the data collected, and the aligned interventions deployed *within a single workshop event*. Each individual thin-slice is uniquely designed to inform the future design of *that workshop* by providing an educational developer with actionable data on participant learning gains.

A short, pre/post assessment on a subset of workshop LOs is not necessarily a new or innovative concept. However, the ultimate goal is to collect enough thin-slices across workshop events to allow one to answer larger-grained questions about one's EdDev practice and impacts, which is what we call the *T-SEDA Process*. Aggregating thin-slices through the T-SEDA Process allows one to investigate impacts across all LOs within a workshop or begin to explore broader impacts across CTL workshops. Either option is useful, but in this paper we primarily focus on the latter. For example, are workshops consistently improving participants' comprehension of evidence-based teaching and learning principles, application skills, or both? If so, what workshop strategies are associated with these gains (or lack thereof)? Over time, the T-SEDA Process is a practical and feasible way to measure the impact of our work on a larger scale and, at the same time, a way to address the challenges of deploying direct, pre/post assessments during each workshop event.

In this paper, we describe how we use thin-slices and the T-SEDA Process to iteratively refine our programming for graduate student instructors and then offer principles and lessons learned for adapting it to other CTL contexts. First, we use a series of case studies to illustrate how thin-slices provide actionable, direct measures of learning gains at the level of individual workshops, across a variety of LOs and EdDev topics. Next, we preview how we are aggregating and comparing across thin-slices to explore the broader impacts of our workshops and the efficacy of common workshop strategies. Additionally, we highlight the key principles underlying the T-SEDA Process and how they can transfer or be adapted to other CTL contexts. Finally, we share the value of our approach for fostering a collaborative culture of formative assessment at CTLs and how this experience has positively impacted our practice as educational developers. While sharing practical strategies and lessons learned, this paper also contributes to a conspicuous gap

in the scholarship of educational development identified by assessment experts (see reviews in Beach et al., 2016; Chism et al., 2012; Haras et al. 2017; Hershock et al., 2022; Stes et al., 2010) by exploring new perspectives and data on the use of direct, pre/post measures of educational development outcomes for formative CTL assessment.

Collecting Thin-Slices to Iteratively Refine Individual Workshops

The Context of Our Workshops and Thin-Slices

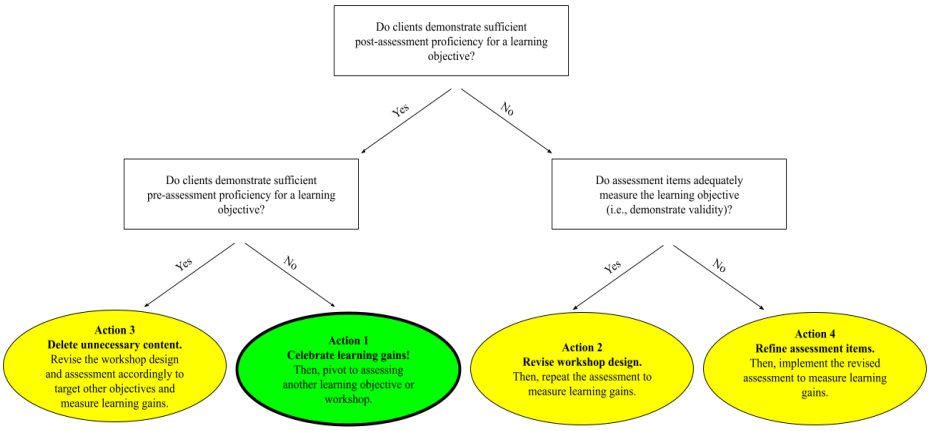
Over four semesters, we piloted the T-SEDA Process across eight 90-minute workshops for graduate student instructors. Workshop topics varied, focusing on multiple aspects of teaching inclusively as well as leading discussions, grading and providing feedback, fostering metacognition, and teaching in recitations or office hours. We offered five of these workshops multiple times, for a grand total of 17 workshop events in which we collected and analyzed a thin-slice (one in fall 2018, five in spring 2019, five in fall 2019, and six in spring 2020). During this time period, our CTL offered an average of 13, 90-minute workshops per semester for graduate students. Workshops occurred as part of our one-day, campus-wide TA orientation or Graduate Student Seminar Series and served 35-200 instructors each, depending on the topic and delivery modality. Trained educational developers designed and facilitated all workshops to achieve three to four LOs, prioritize active learning, apply to all participants' disciplines and teaching contexts, disseminate research-based learning principles and associated, practical teaching strategies, highlight effective technology-enhanced learning strategies, and enhance diversity, equity, and inclusion in teaching. Attendance was optional, but may have been required by academic departments for TA training.

Embrace Four Possible Actions Resulting from a Thin-Slice

From this process, we have identified four possible actions to take after collecting a thin-slice (Figure 2). These actions are not mutually exclusive. Which "actions" we take depends on two things. First, we consider the results of the thin-slice, both the observed gains and post-assessment performance. Second, we consider a number of contextual factors, which may vary (e.g., relative priority of the learning objective targeted, the relative strength of the EdDev intervention, potential return on investment (ROI) of further iterating workshop elements or assessments, available resources, etc.

Figure 2

Decision Tree for Interpreting Actionable Thin-Slice Outcomes



When data shows that participants do not demonstrate proficiency on pre-assessments, but we observe learning gains on post-assessments, then we celebrate the impacts of the workshop and consider the opportunity to *pivot* our focus *to assessing another LO or workshop* (Action 1). We do not have a definitive threshold for what constitutes a sufficient learning gain. Instead, judgments depend on the contextual factors enumerated above as well as the difficulty of the LO, especially LOs requiring higher-order skills. Adhering to an arbitrary performance threshold alone would likely be counter productive for determining ROI on continued iteration. Furthermore, 100% post-assessment proficiency is not the only appropriate marker of meeting program goals, given the context of the intervention (e.g., the impact of a 30-minute intervention during a 90-minute workshop, the lack of stakes for participants, measurement constraints, and limited CTL resources).

When participants approach maximum proficiency for a particular LO on the pre-assessment, then Action 2 is to *cut unnecessary content*. We want programming to align to the needs of our instructors, given limited time and resources. This result signals the opportunity to enhance ROI for participants and us by eliminating content for which attendees are proficient and then increasing either depth or breadth.

In contrast, when performance on pre-assessments is low and learning gains are minimal, Action 3 is to *revise the workshop design*. To improve learning gains, we may revisit the content itself, time on task for LOs, the nature of practice and feedback during the workshop, or the alignment of the design with our LOs.

When analyses suggest the assessment items are not optimally measuring outcomes (e.g., an item is poorly designed or too difficult given time allotted), Action 4 is to *refine the assessment items* to improve the validity of the measurement and resulting inferences.

Again, the above actions are not mutually exclusive. A thin-slice may target two LOs and data may suggest the same or different action items for each LO (e.g., revising the assessment items *and* cutting content, respectively). In any event, we decide which action(s) to take on a case-by-case basis, considering not only the thin-slice data, but also judging which action(s) will maximize ROI on formative assessment and program (re)design efforts moving forward. The four examples of thin-slices below illustrate these action items.

Examples of Thin-Slices and Resulting Actions

For each case study below, we summarize the thin-slice LO(s), assessment items, findings, and implications for our practice. Across a variety of workshop designs, objectives, and sample sizes, these cases illustrate each of the four, non-mutually exclusive actions that can result from a thin-slice (Figure 2). Case 1 illustrates how an authentic assessment task measured learning gains and validated the efficacy of associated workshop elements. Case 2 demonstrates how both the design and analysis of the thin-slice informed targeted workshop revisions, including the removal of content for which participants demonstrated initial proficiency. Case 3 illustrates the evolution of thin-slice assessment items to iteratively improve measurements and the use of counterbalanced multiple choice questions. Case 4 highlights the value of qualitative analyses to document learning gains as well as inform workshop and assessment refinements.

Case 1: A Thin-Slice Can Validate a Workshop Design by Directly Measuring Learning Gains

A high priority objective of our workshop on grading and feedback is to train instructors to effectively yet efficiently provide feedback on student

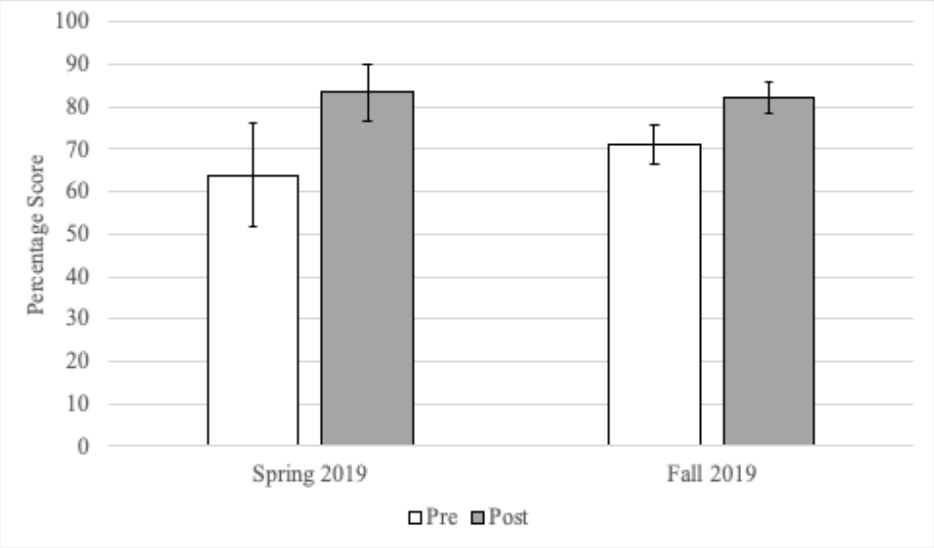
work, a skill aligned with TA responsibilities. During the workshop, participants practiced this skill authentically during an activity representing approximately 25% of workshop content. Thin-slice pre/post tests used a similar authentic task. Participants reviewed a quantitative prompt question, sample student response, three-criteria rubric, and student's rubric score. We tasked respondents with generating effective feedback for the sample student. We evaluated responses quantitatively, using a trait-analytic rubric based on the characteristics of effective feedback taught in the workshop (Ambrose et al., 2010). As for all thin-slices, we de-identified samples, such that coders did not know whether they were evaluating a pre- or post-assessment. A Wilcoxon Signed Ranks Test assessed the difference between pre- and post-assessment scores. In spring 2019, participants scored significantly higher on the post-test (M = 83%) compared to the pre-test (M = 64%), 30 respondents, $Z = 2.70$, $p < .01$). Respondents improved by more clearly signaling what is correct (and not) and providing feedback that better balances positive and constructive elements. After replicating results in fall 2019 (post-test M = 82%; pre-test M = 72%, 163 respondents, $Z = 4.07$, $p < .001$; Figure 3), and considering all contextual factors (see above), we considered shifting the focus of future thin-slices to other LOs and workshops (Figure 2). This case illustrates how a thin-slice can help you determine how best to allocate limited resources for future assessment and workshop design efforts.

Case 2: Thin-Slice Assessment Designs and Data Each Inform Workshop Revisions

Our workshop on designing and facilitating discussions aims to teach strategies for actively engaging students and designing effective discussion questions. While retrofitting a thin-slice to this workshop, we discovered an alignment gap. Consequently, we redesigned the workshop to include more resources and practice for crafting discussion questions before collecting any data.

The eventual assessment leveraged counterbalanced vignettes followed by an authentic lesson planning task. Vignettes named a topic and pre-work assigned to students. For example, "This week's topic is tariffs and trade wars. Students are reading an op-ed on the recent US tariffs on Chinese imports and a book chapter on the World Trade Organization. Please discuss these readings with students during recitation." Respondents then com-

Figure 3
Participants' Scores on an Authentic Task
Before and After the Quantitative Grading and Feedback Workshop



Note: Error bars are 95% confidence intervals for means.

pleted this authentic task: (1) describe a discussion facilitation strategy, and (2) write one discussion question that would help students practice evaluating different perspectives. We quantitatively scored responses using a trait-analytic rubric based on the principles the workshop emphasized.

Across two workshop offerings, Wilcoxon Signed Ranks Tests revealed that participants exhibited different starting proficiency and gains, but similar final levels of proficiency (spring 2019, 22 respondents, pre-workshop 57%, post-workshop 75%, $Z=1.10$, $p=.27$; spring 2020, 49 respondents, pre-workshop 70%, post-workshop 78%, $Z=1.99$, $p<0.05$). After each workshop, participants exhibited relatively high proficiency, describing more structured, active learning techniques likely to engage most or all students simultaneously. In contrast, pre-assessments demonstrated relatively high proficiency for designing discussion questions (>85% aligned questions with LOs and their choice of active learning strategy; 73% crafted open- rather than closed-ended questions). Because many attendees entered with high profi-

ciency, we had to consider whether to: (1) reprioritize LOs; (2) eliminate unnecessary content on designing discussion questions; and/or (3) add content to target advanced skill development. Ultimately, we implemented options 2 and 3. This case illustrates how intuition and participant feedback alone are insufficient to guide workshop development. The systematic thin-slice design process combined with data mitigated that blindspot.

Case 3: Thin-Slice Data Informs Refinement of Assessments

Our workshop on working 1-on-1 with students targeted two LOs: (1) identify developmental stages of students, and (2) describe strategies to support students' intellectual development. The workshop contained a presentation of longitudinal research on undergraduate student development (Baxter-Magolda, 1992). Participants then practiced identifying the developmental stage based on student quotes and worked in pairs to brainstorm strategies to support each student's development. Similarly, the open-ended assessment items prompted participants to: (1) read a student quote and identify their intellectual development stage, and (2) describe a strategy to support development to the next stage.

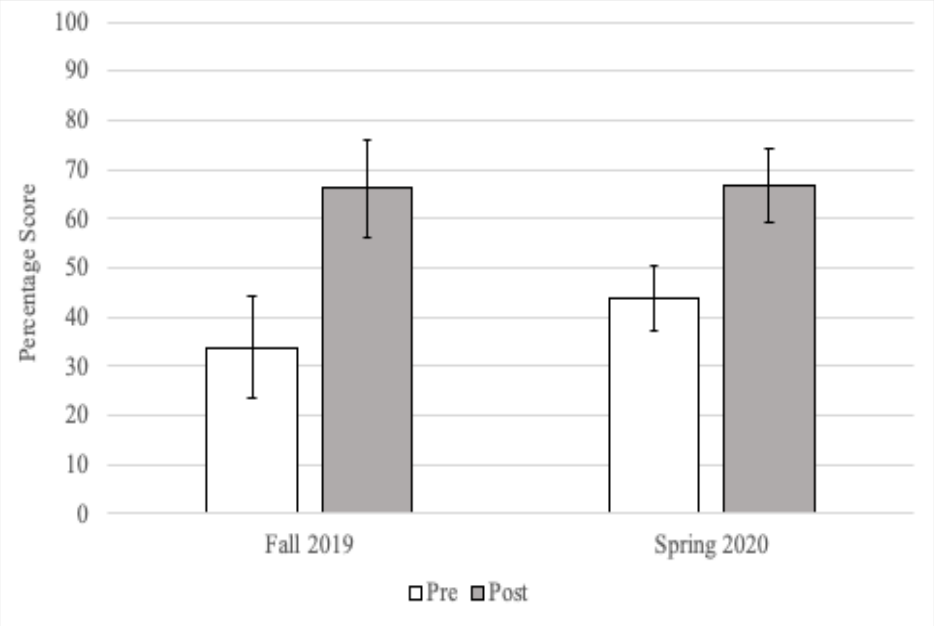
Our pilot assessment indicated that this open-ended approach did not elicit sufficient evidence of knowledge. On the post-test, many participants simply named intellectual development stages without defining them (Appendix B.1) and did not sufficiently describe supporting strategies. Results, therefore, were inconclusive.

For our second attempt, we tried multiple choice questions (MCQs). We presented participants with a case study including the same student quotes but replaced open-ended questions with two MCQs. Unfortunately, the data suggested that we constructed cognitively challenging, double-barreled answer choices. Each answer choice described a stage of intellectual development *and* a possible teaching strategy (see Appendix B.2). Based on the data, we could not determine whether: (1) the workshop design did not impact learning, or (2) the complexity of the answer choices confounded the data.

In our third and fourth iterations, we parsed each MCQ into two, asking separately about (1) students' developmental stage, and (2) a supporting strategy (Appendix B.3). This design addressed previous challenges. Results from two semesters suggested positive learning outcomes on both LOs (Figure 4). Wilcoxon Signed Ranks Tests tested for significant differences between pre- and post-assessment scores. In fall 2019, respondents scored significantly higher on the post-assessment ($M = 66.18$, $SD = 28.79$) compared to

the pre-assessment ($M = 33.82$, $SD = 29.44$, 35 respondents, $Z = 3.49$, $p < .001$). In spring 2020, respondents also scored significantly higher on the post-assessment ($M = 66.78$, $SD = 31.47$) compared to the pre-assessment ($M = 43.84$, $SD = 28.80$, 82 respondents, $Z = 3.94$, $p < .001$). Based on this validation of workshop efficacy, we next considered pivoting to assess other LOs. Overall, this case illustrates the thin-slice action (and added value) of iteratively refining assessment items to provide actionable data. Specifically, we learned transferable lessons about how best to adapt MCQs as efficacious, efficient alternatives to open-ended thin-slice items.

Figure 4
Respondents' Scores Before and After Revising Thin-Slice Assessment Items in a Workshop on Working Well One-on-One



Case 4: Qualitative Analyses Add Value to Thin-Slices

In a new workshop on how to center diversity, equity, and inclusion (DEI) in course design, we emphasized two questions: (1) How do DEI issues affect student learning?; (2) How can I center DEI in my LOs, course content, and assessments? The workshop leverages active learning exercises to explore each question. The thin-slice included two open-ended questions: "Why is it important to center DEI in course design?" and "What are ways that instructors can center DEI in course design?" Due to a small sample (13 respondents) and the potential to demonstrate learning gains in myriad, nuanced ways, we did not quantitatively analyze responses. Instead, we read participants' responses holistically, qualitatively coding them based on the presence of key concepts, terminology, and strategies highlighted during the workshop.

When asked about the importance of centering DEI in course design, before the workshop, many participants provided general rationales such as the importance of students feeling included or "to account for a plurality of perspectives and ways of learning." After the workshop, participants also explained why DEI was important in terms of its impact on specific student outcomes. For example, one participant wrote, "It helps students feel they belong, which increases their motivation, which improves the student's learning outcomes." This response represents a conspicuous trend following the workshop. By linking DEI to specific, educational outcomes and terminology (e.g., student belonging, motivation, learning outcomes), participants signaled a more concrete understanding of the mechanisms by which DEI supports student success.

Additionally, after the workshop participants named more ways instructors can center DEI in course design, whereas before, they only mentioned altering course content. Afterwards, participants also identified LOs and assessments as opportunities to integrate DEI. One stated, "At each stage of course design …, instructors can incorporate DEI: adjust or extend LOs to incorporate DEI, make sure course content is inclusive of diverse voices, make sure assessments are inclusive/take into account DEI. Consider different kinds of diversity: racial, language, ethnic, disability, gender."

While participants made progress towards achieving LOs, trends were not ubiquitous. Nonetheless, qualitative analyses identified two conspicuous, actionable enhancements to explore in future iterations. Although respondents broadened their perspective on ways to center DEI in course designs, most did not articulate specific strategies, a result we cannot explain.

Future iterations could adjust workshop content and strategies and/or assessment items to better understand learning gains (or lack thereof).

Aggregating Thin-Slices to Infrom EdDev Practice Across Workshops

So far, we've discussed single thin-slices; however, in this section, we will demonstrate how the T-SEDA Process can aggregate results across thin-slices to further inform workshop design and strategic planning. For example, an important question to both our EdDev practice as workshop designers and our CTL's strategic planning is: *Are instructor learning gains from workshops limited to low-level cognitive skills (e.g., comprehension) or can workshops also enhance higher order cognitive skills (e.g., analysis, application)?* If workshops are not developing our participants' higher-order cognitive skills regarding teaching, then we need to consider alternative, less scalable strategies for group programs to achieve those LOs. To date, the Scholarship of Educational (SoEd) literature does not include studies of the above question. Below, we present preliminary results demonstrating how a CTL could use the T-SEDA Process to answer this question by aggregating thin-slices across four semesters (fall 2018-spring 2020) and 17 workshop events. Because each thin-slice contained 1-2 LOs, we compared analyses across thin-slices at the grain-size of an individual LO. Overall, these 17 thin-slices target 15 unique LOs and include 21 independent, direct measurements of those LOs (n=5 comprehension LOs; n=16 analysis or application LOs). We do not present a formal quantitative meta-analysis of learning gains due to our small sample size and because some thin-slices leveraged qualitative research methods. Instead, we describe preliminary observed patterns and their implications.

Across thin-slices, our data suggest that our approach to 90-minute EdDev workshops fosters instructor learning gains for both lower- and higher-order cognitive skills. Four of five measurements (80%) showed significant learning gains for lower-order LOs targeting comprehension. Eight of sixteen measurements (50%) demonstrated significant learning gains for higher-order LOs targeting analysis or application. Additionally, we intentionally repeated thin-slices containing three comprehension LOs and four analysis/application LOs. In all seven cases (100%), we replicated learning gains from a previous thin-slice. Our preliminary data suggest that achieving higher-order LOs in a 90-minute workshop is harder than fostering comprehension regarding evidence-based teaching and learning principles.

However, our data also suggest that designing workshops to foster higher-order instructor learning gains is certainly feasible.

These preliminary results illustrate the added value of the T-SEDA Process for evidence-based EdDev practice. First, it allowed us to explore unanswered questions regarding the broader impacts of a key element of our CTL's programming when the SoED literature did not provide guidance. Second, the T-SEDA Process informed our CTL's strategic planning for programs, services, and resource allocation. Given the staff effort our CTL invests in workshops, these direct outcomes measures validated the ROI.

Third, when workshops did not achieve LOs on the first attempt, the T-SEDA Process empowered us to iteratively refine our designs, informed by direct measures of impacts, rather than intuition or participant feedback alone. Seven of 21 times, the T-SEDA Process data suggested that we needed to redesign a workshop to better achieve LOs or meet instructors' needs. Only twice did thin-slices generate completely uninterpretable and unactionable data due to suboptimal assessment item construction. In all cases, we refined the workshop design and/or assessment instrument and ultimately achieved 12 of 15 (80%) targeted unique LOs. Because we are leveraging direct measures of impacts, rather than intuition and participant feedback alone, we are confident that the T-SEDA Process will enable us to continue to refine our workshops and achieve the remaining three LOs.

Fourth, the T-SEDA Process continues to inform our general approach to workshop design in practical, concrete ways. For example, in 17 of our 21 direct measurements, the workshop interventions aligned with LOs represented only 15-35 minutes of a 90-minute workshop's content. Consistent with research on active learning (Freeman et al., 2014), these results suggest that relatively small interventions can support substantive learning gains. Moreover, these preliminary results suggest that workshop designs prioritizing 2-3 LOs total are reasonable in scope for 90-minute events.

Later in this paper, we also highlight the added value of the T-SEDA Process for fostering a collaborative CTL culture of formative assessment and positively impacting the professional development of individual educational developers. Below, we devote entire sections to each topic, given their complexity and nuance. However, first we discuss how the T-SEDA Process is transferable to other EdDev contexts.

Transferring the T-SEDA Process to Other CTL Contexts

To this point, we have described our adoption of and the early results from the T-SEDA Process. In this section, we offer additional ideas and encourage CTL's to expand and adapt the T-SEDA Process to fit their context and programming.

Is the T-SEDA Process Limited to Workshops?

No. The T-SEDA Process can be applied to other CTL group programs, such as orientations, institutes, and learning communities, regardless of modality (i.e., in-person or remote, synchronous or asynchronous). For instance, an authentic, comprehensive pre/post assessment for a week-long course design institute could be quite long or impractical given logistical constraints or limited CTL assessment resources. However, the T-SEDA Process provides a feasible, actionable approach to formative assessment, within or across particular days/segments of the institute. Moreover, when paired with other direct measures of impacts, such as analyses of participants' deliverables (e.g., syllabi from before and after the program), the T-SEDA Process provides desirable, complementary data.

Is the T-SEDA Process Feasible for Small CTLs or When Resources Are Limited?

Yes. A CTL can adjust the scope and staffing of the T-SEDA Process to meet their needs and context. For comparison, our CTL employs both teaching and assessment consultants (six and two, respectively, during the pilot described above). Currently, we typically assign one of each consultant type per thin-slice to leverage their complementary skill sets. However, we believe a single educational developer could design, implement, and analyze a thin-slice within one semester. While collaborating with an assessment expert is certainly an advantage, teaching consultants possess many transferable skills. For instance, designing a thin-slice uses many common skills across educational developers, including backwards design (of workshops), articulating student-centered and measurable LOs, and alignment. In addition, they frequently support instructors by helping them design effective rubrics and course assessments, including mechanically sound exam questions. These assessment skills transfer directly to the design and analysis of thin-slices that can yield actionable data.

If human resources are limited, within CTLs, we also recommend surveying staff for relevant assessment skill sets. We discovered that several of our teaching consultant colleagues possessed deep expertise with qualitative research methods, complementing the expertise of our assessment consultants. This discovery enabled thin-slice approaches that did not depend on quantitative analytical methods or associated expertise. Alternatively, one could form networks of support with colleagues across institutions, especially during the thin-slice design phase and the aggregation of thin-slices during the T-SEDA Process. Such partnerships are also valuable for sharing assessment instruments and lessons learned as well as preventing duplication of effort.

Additionally, we now deploy thin-slices strategically, based on available resources. In other words, we only deploy thin-slices in a subset of workshops each semester, based on staff capacity and strategic prioritization of formative assessment needs, rather than applying the T-SEDA Process across every possible group program. Specifically, we use two diagnostic questions to prioritize and evaluate capacity. What do we wish we knew about our participants' learning? And, what would we do with the data? If the answer to the second question is "not much," then we deprioritize that potential thin-slice. In some situations, it may not be worth the overhead to collect and analyze thin-slice data. For example, rechecking a previously assessed workshop for quality assurance may be less valuable than focusing efforts on a new offering or different LO.

Similarly, efforts can be focused disproportionately during the more amenable times of year. For example, moving forward, we plan to allocate comparatively more effort to collecting thin-slices in spring than fall semesters, when we have more turnaround time for analyses during summer months. We also evaluate when to pause during our T-SEDA Process and reallocate resources towards other LOs, programs, or assessment strategies (e.g., during the intensive pivot to remote instruction during the 2020 pandemic). These decisions often depend on contextual factors as well as the possible outcomes within the T-SEDA Process (Figure 2).

The approaches above allow us to manage our limited resources strategically and sustainably. Thin-slices are not meant to assess everything. Assessing only one group event per semester does generate actionable data. And, as long as thin-slices are implemented regularly over time, data will gradually accrue through the T-SEDA Process. Instead of thinking, "only large CTLs have the resources for the T-SEDA Process," we suggest starting small, targeting fewer workshops and LOs per semester, including fewer assessment items per thin-slice, and prioritizing a patient, formative lens.

In our experience, the return on investment from the T-SEDA Process is worth the time and resource costs. We always learn something formative, whether it is about the workshop, participants, assessment, or our EdDev practice. In the next section, we enumerate the guiding principles for adapting the T-SEDA Process that emerge from case studies above.

Designing Thin-Slices and Implementing the T-Seda Process: Guiding Principles

1. Apply Backwards Design Frameworks to Workshop Development

When designing workshops from scratch, first articulate the LOs. Next, design the thin-slice's pre/post assessment to align with LOs, *before* designing workshop content. This backwards design (Wiggins & McTigue, 2005) sequence prioritizes alignment and informs assessment choices. We intentionally design our workshop learning objectives, assessments, and evaluation criteria to align with appropriate evidence-based principles of learning (e.g., Ambrose et al., 2010; Benassi et al., 2014). In our experience, designing the assessment last is more difficult (but not impossible) and more likely to result in poor alignment and less informative measurements.

2. Adopt a Long-View, Rotating the LOs Assessed as Workshops Are Repeated

Limit the scope of each thin-slice assessment. Focus on one to two LOs, prioritizing what would be most informative for the next workshop iteration. By constraining the scope to a "thin slice" of each workshop, we mitigate time constraints without sacrificing workshop content. Over time, the focal LOs can shift within a workshop. As the foci of thin-slices rotate over time, data accumulates, painting a more complete picture of impacts across LOs. This long view is particularly amenable to holistically assessing individual workshops, but can also be applied across workshops to explore larger-grained impacts (i.e., the T-SEDA Process, see above).

3. Leverage Existing Workshop Activities as Assessment Opportunities

Programs often include "warm up" activities to jumpstart engagement, establish active learning norms, and/or activate prior knowledge. Consider

whether these exercises can function as the pre-assessment. Similarly, workshop exercises where participants practice applying principles to authentic tasks can be leveraged as post-assessment opportunities, as long as they occur after aligned instruction. In Cases 1-3 above, for example, several of the pre/post assessment items are modified versions of embedded workshop activities. Alternatively, thin-slices can leverage deliverables generated through Classroom Assessment Techniques (Angelo & Cross, 1993) embedded in workshops. These approaches capitalize on creative replacement of, rather than addition to, workshop content.

4. Remember Less Can Be More

As standard practice, we ask ourselves, "What component of the workshop could we eliminate spontaneously, if needed?" This question forces us to prioritize and embrace that less (*content coverage*) can be more (*for learning*). By cutting lower priority content, we aim to create space for valuable formative assessment opportunities. When implementing a thin-slice, we also repurpose time previously allocated to feedback surveys to reduce the challenge of cutting content. Because we also value participant feedback, we deploy feedback surveys in workshops that are not currently employing thin-slices.

5. Keep Assessments Short

Thin-slices leverage pre- and post-assessments that can be completed in five minutes. Instruments contain two to four questions or one to two brief authentic tasks measuring knowledge or skills (see Appendices A and B). Examples of brief authentic tasks include writing a discussion question (see Case 2) or student-centered and measurable LO, crafting a multiple choice exam question, or redesigning a slide to be more effective for learning. Having more assessment items aligned with an LO may increase the probability of a valid measurement, but is not necessarily required for effective formative assessment.

Furthermore, we recommend applying basic survey and assessment item design principles to thin-slice designs. For example, provide sufficient direction to ensure that individuals from all disciplines can engage. Avoid complicated questions that ask about multiple constructs. Separating items reduces the likelihood that respondents miss instructions. And, reduce cognitive load by simplifying an item's contextual, surface information or requested deliverable.

6. Counterbalance Isomorphic Questions, When Possible

Counterbalancing helps determine whether observed learning gains are caused by the workshop or by assessment item construction. We counterbalance by using isomorphic versions of assessment items on pre- and post-assessments (Appendix A). Half of the respondents receive pre-assessment version A and post-assessment version B, and vice versa for the other respondents. Comparing results for counterbalanced items is particularly helpful when exploring new assessment strategies. This strategy was used for Cases 1, 2, and 3 above.

7. Authentic Tasks Can Be Assessed with Quantitative and/or Qualitative Analyses

Quantitative and qualitative methods each have pros and cons. Together, they provide complementary, rigorous tools for actionable, formative assessment. Quantitatively knowing *the magnitude of change* or *how many participants changed* (or not) is informative. However, direct measures of learning are not limited to numbers and quantitative analyses. Qualitatively understanding the nuances of change is often equally important (see Case 4 above). While a mixed methods approach may be optimal, given limited resources and other contextual factors, either analytical approach alone can still be a viable option for a thin-slice. Regardless, we recommend challenging one's disciplinary biases regarding qualitative and quantitative methods to avoid limiting one's options.

8. Avoid Implicit Bias by De-Identifying Thin-Slice Data Prior to Analysis

We de-identify data in two ways. First, names or other identifiers with the potential to bias coders are removed from pre- and post-assessments prior to analysis. For in-person seminars, handouts contain an identifying number that is used to pair pre- and post-assessments. Additionally, prior to analysis, individual responses are de-identified regarding whether they came from a pre- or post-assessment.

9. The Absence of a Comparison Group Is Not a Deal-Breaker

Historically, research on EdDev impacts lacks comparison groups (Chism et al., 2012; Stes et al., 2010). Without comparison groups, pre/post assessments document "what happens," but not necessarily "what causes." Two recent studies leverage institutional data to generate viable comparison groups (Hershock et al., 2022; Meizlish et al., 2018). For a thin-slice, an ideal comparison group includes instructors who did not use the service, but are otherwise comparable. Such comparison groups may not be feasible, particularly for stand-alone workshops. Nevertheless, even without a comparison group, the case studies above illustrate how the T-SEDA Process can provide actionable, formative data. While one must interpret data cautiously regarding causality, it is unlikely that immediate learning gains (with strong effect sizes) observed during a 90-minute workshop would be caused by something other than the workshop itself.

10. Be Transparent

We developed a script that our educational developers use to clearly communicate at the beginning of workshops what we are doing and why, given our core values regarding data-informed practice. We emphasize that we use the anonymous data to improve our services, rather than summatively evaluate participants or facilitators. Attendees may opt out for any reason.

Fostering an Inclusive CTL Culture of Formative Assessment

Adopting the T-SEDA Process represents a tangible shift in EdDev practice and culture. Cultivating a healthy culture of formative assessment requires CTL leadership and staff to play important roles. The ideal transition is carefully framed and managed collaboratively, inclusively, and transparently.

Our CTL gradually integrated the T-SEDA Process into our standard toolkit for assessing workshops starting in fall 2018. Four semesters of experience reinforced the value of several strategies we employed to foster and sustain an inclusive CTL culture of formative assessment: *focus on formative implementation*; *transition slowly and start small*; *promote collaboration and community ownership*; *continuously solicit staff feedback and adjust workflows and expectations*; *periodically debrief results and lessons learned*; and *maintain institutional memory and documentation*. Below, we share reflections and perspec-

tives from our CTL's staff and leadership that represent our general consensus rather than ubiquitous agreement.

Focus on Formative Implementation

CTL leaders might be tempted to use the T-SEDA Process for summative staff performance evaluations. Don't, even if you use this approach to demonstrate value to external stakeholders. The T-SEDA Process is designed to focus on workshop impact, not an individual educational developer's impact or performance, which is influenced by many competencies that are not measured directly by thin-slices (e.g., facilitation skills). When our leadership proposed the T-SEDA Process to staff, some colleagues expressed apprehension regarding its potential for summative performance evaluations. The T-SEDA Process as an evaluative tool is anathema to collaborative, formative assessment. Instead, communicate that null results from thin-slices are valuable formative opportunities to learn and iterate on a service design. Reassure staff that thin-slices will *never* be used to summatively evaluate for merit and promotion. For individual staff, the benefits of this approach center professional development opportunities, including collaboration on service design, establishing clear and measurable learning objectives, and designing well-aligned EdDev experiences. Before implementing the T-SEDA Process, we recommend transparently setting these expectations.

Transition Slowly, Start Small

We intentionally piloted the T-SEDA Process slowly in a manageable subset of our workshops, with multiple onboarding sessions for discussing the underlying principles and philosophy, answering questions, preemptively exploring hypothetical scenarios, and soliciting feedback on implementation. Over time, we gradually increased the breadth of workshops implementing thin-slices, carefully considering staff input on feasibility. At present, staff even proactively request including a thin-slice or propose creative ideas for new assessment items.

Promote Collaboration and Community Ownership

Certainly, an individual educational developer could design and implement a thin-slice. However, colleagues can contribute diverse, complemen-

tary perspectives and skill sets while collaborating. Furthermore, a community-based approach fosters shared, rather than proprietary, workshop ownership. We continue to use a scaffolded approach as staff learn to integrate the T-SEDA Process into their practice. Encouraging shared ownership and professional development is a key part of the process.

Periodically Debrief Results and Lessons Learned

We discovered that without regular debriefs, lessons learned remain siloed. Community debriefs, within and across thin-slices, invite staff feedback and perspectives critical to refinement. They also promote collaboration, collective ownership, and professional development. Conducting regular debrief meetings sustains a growing culture of formative assessment.

Continuously Solicit Staff Feedback and Adjust Workflows and Expectations

Repeatedly soliciting staff feedback can help lower barriers to adoption, manage workloads humanely, increase efficiency, and maximize return on investment. Consistent feedback not only refines processes but also fosters an inclusive climate during the four phases of implementing the T-SEDA Process: design, deployment, data analysis, and application of results to future practice. We found adopting the T-SEDA Process without adjusting other strategic priorities overloaded staff. Similarly, as competing demands change, we periodically adjust our workflows, effort, and expectations.

Maintain Institutional Memory and Documentation

Taking time to plan and build infrastructure for institutional memory will pay dividends, especially given staff turnover and rotating workshop assignments. We proactively created an accessible, structured, and user-friendly archive of workshop designs, thin-slice materials, data, and recommended future changes. The purpose of the archive is to increase efficiency when an educational developer inherits an existing program and to sustain data-informed iterations of our programs. Without this archive, lessons learned are more likely to be lost, even with regular debrief meetings.

Looking Forward: How Can the T-SEDA Process Positively Impact Educational Developers?

In this article, we explored how to directly measure instructors' learning gains from EdDev workshops and foster a collaborative culture of formative CTL assessment. We learned many practical lessons about the T-SEDA Process, including that it is customizable, feasible, and sustainable given conspicuous logistical constraints and resource limitations, and can be implemented at scale. More importantly, we experienced actionable results and gained valuable insights into our Ed Dev practices within and across thin-slices at each stage of the process.

After adopting the T-SEDA Process, we observed the following positive impacts on our workshop design and consulting practices. Designing thin-slices has challenged our assumptions about what is important and realistically achievable during a 90-minute workshop, proving an effective lens for prioritizing our LOs and surfacing gaps that prompt revision. Additionally, when we debrief workshops together, we draw on the direct measures of how particular designs or facilitation strategies impact outcomes. Consequently, discussions are more conspicuously evidence-based rather than relying on intuition and participant feedback. The T-SEDA Process also enhanced our collective fluency regarding basic principles of measuring outcomes and the mechanics of designing effective assessment items and rubrics. Furthermore, when consulting with instructors, we are more likely to take "an assessment perspective" because habits of mind from the T-SEDA Process and designing thin-slices transfer. For example, we look for targeted opportunities to collect and leverage baseline data regarding student outcomes. This baseline data, analogous to the pre-assessment in a thin-slice, helps target where to focus efforts to improve teaching and learning. Then, when helping an instructor implement an intervention to enhance student learning, we consider and discuss how to determine its efficacy, including identification of direct data sources already embedded in the course design. Overall, in our experience, piloting the T-SEDA Process has not only fostered a collaborative culture of evidence-based EdDev practice at our CTL, but also contributed conspicuously to our individual professional development as consultants and program designers.

Educational developers and CTLs can also benefit from the T-SEDA Process more broadly. First, the T-SEDA Process can help answer the need for more direct measures of EdDev outcomes (Beach et al., 2016; Gilmore & Hatcher, 2021), leveraging one potential data source, instructor learning

gains. Second, if an institution's administration asks for evidence of a CTL's impact, the T-SEDA Process is a feasible option to consider *if presenting data in the aggregate across workshops.* We do not recommend summative assessments of individual CTL workshops (see the cautions above regarding potential negative impacts on CTL culture). Ideally, strategic plans for data-driven outcomes assessments would leverage multiple data sources (Ellis et al., 2020; POD Network, 2018; Wheeler & Bach, 2020), but the T-SEDA Process represents a practical and potentially complementary strategy to add to a CTL's assessment toolkit. Third, through the T-SEDA Process, educational developers can visibly model evidence-based practice. Data-driven decision making is lauded at our research-intensive institution. And, our CTL's mission includes facilitating evidence-based teaching by distilling and disseminating education research into practical pedagogical strategies across teaching contexts, as well as supporting instructors and programs as they collect, interpret, and act upon student outcomes data (Lovett & Hershock, 2020). By transparently embracing the T-SEDA Process, we are demonstrating evidence-based practice to our participants and stakeholders, analogous to what we advocate and actively support at the course- and program-level.

Our CTL will continue to implement the T-SEDA Process as part of our formative assessment toolkit. We hope this paper will inspire creative adoption of the T-SEDA Process in different EdDev contexts *and* more scholarship on the impacts of EdDev. In particular, we hope future research will investigate the long-term retention of instructor learning gains from workshops and how they translate to changes in teaching practices and student outcomes. As our thin-slice database grows and formative assessment practices evolve, we also hope to contribute our findings to filling this gap in the EdDev literature.

Acknowledgements

It's all Marsha Lovett's fault. In 2018, she challenged us to "raise our game" regarding direct, formative outcomes assessments of our CTL's services. The T-SEDA Process resulted from this challenge. The Institutional Review Board (STUDY2016_00000148) approved the use of our data from T-SEDAs as research at the university where they were conducted. We thank Katherine Phelps Walsh and two anonymous reviewers for their constructive feedback on earlier versions of this paper. Judy Brooks designed Figure 1.

References

Ambrose, S. A., Bridges, M. W., DiPietro, M., Lovett, M. C., & Norman, M. K. (2010). *How learning works: Seven research-based principles for smart teaching*. Jossey-Bass.

Angelo, T.A., & Cross, K.P. (1993). *Classroom assessment techniques: A handbook for college teachers* (2nd Ed.). Jossey-Bass.

Barrett, M., Hershock, C., McCarthy, M., Melville, M., & Mertz, J. (2021). What type of debrief is best for learning during think-pair-shares? *Teaching & Learning Inquiry, 9*(1), 45–60.

Baxter-Magolda, M. (1992). *Knowing and reasoning in college: Gender-related patterns in students' intellectual development*. Jossey-Bass.

Beach, A. L., Sorcinelli, M. D., Austin, A. E., & Rivard, J. K. (2016). *Faculty development in the age of evidence: Current practices, future imperatives.* Stylus Publishing.

Benassi, V.A., Overson, C. E., & Hakala, C. M. (2014). Applying science of learning in education: Infusing psychological science into the curriculum. Retrieved from the Society for the Teaching of Psychology website: http://teachpsych.org/ebooks/asle2014/index.php

Chism, N.V. N., Holley, M., & Harris, C. J. (2012). Researching the impact of educational development: Basis for informed practice. *To Improve the Academy,* 31(1), 129–145.

Cook, C. E., & Kaplan, M. (Eds.). (2011). *Advancing the culture of teaching on campus: How a teaching center can make a difference.* Stylus Publishing.

Eberly Center (2021, February 16). *Teaching as Research: Discovering what works for student learning.* https://www.cmu.edu/teaching/teaching-as-research/index.html

Ebert-May, D., Derting, T. L., Hodder, J., Momsen, J. L., Long, T. M, and Jardeleza, S. E. (2011). What we say is not what we do: Effective evaluation of faculty professional development programs. *BioScience, 61*(7), 550-558.

Ellis, D. E., Brown, V. M., & Tse, C. T. (2020). Comprehensive assessment for teaching and learning centres: A field-tested planning model. *International Journal for Academic Development*, *25*(4), 337-349.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the national academy of sciences, 111*(23), 8410-8415.

Gilmore, J. and Hatcher, M., (2021). *Preparing for College and University Teaching: Competencies for Graduate and Professional Students.* Stylus Publishing.

Gonyea, R. M. (2005). Self-reported data in institutional research: Review and recommendations. *New Directions for Institutional Research, 2005*(127), 73-89.

Haras, C., Taylor, S. C., Sorcinelli, M. D., & von Hoene, L. (Eds.). (2017). *Institutional commitment to teaching excellence: Assessing the impacts and outcomes of faculty development.* American Council on Education

Hershock, C., Cook, C. E., Wright, M. C., & O'Neal, C. (2011). Action research for instructional improvement. In C. E. Cook & M. Kaplan (Eds.), *Advancing the culture of teaching on campus: How a teaching center can make a difference* (pp. 167-182). Stylus Publishing.

Hershock, C., Pottmeyer, L. O., Harrell, J., le Blanc, S., Rodriguez, M., Stimson, J., Walsh, K. P., & Weiss, E. D. (2022). Data-driven iterative refinements to educational development services: Directly measuring the impacts of consultations on course and syllabus design. *To Improve the Academy: A Journal of Educational Development, 41*(2): 9. doi: https://doi.org/10.3998/tia.926

Hines, S. R. (2015). Setting the groundwork for quality faculty development evaluation: A five-step approach. *Journal of Faculty Development*, *29*(1), 5-12.

Lovett, M., & Hershock, C. (2020). Cultivating and sustaining a faculty culture of data-driven teaching and learning: A systems approach. *To Improve the Academy*, *39*(1), 63-93.

Meizlish, D. S., Wright, M. C., Howard, J., & Kaplan, M. L. (2018). Measuring the impact of a new faculty program using institutional data. *International Journal for Academic Development, 23*(2), 72-85.

POD Network. (2018). Defining what matters: Guidelines for comprehensive center for teaching and learning (CTL) evaluation. POD Network. https://podnetwork.org/content/uploads/POD_CTL_Evaluation_Guidelines__2018_.pdf

Stes, A., Min-Leliveld, M., Gijbels, D., & Van Petegem, P. (2010). The impact of instructional development in higher education: The state-of-the-art of the research. *Educational Research Review, 5*(1), 25-49.

Wheeler, L. B., & Bach, D. (2021). Understanding the impact of educational development interventions on classroom instruction and student success. *International Journal for Academic Development*, *26*(1), 24-40.

Wiggins, G., & McTighe, J. (2005). *Understanding by design* (2nd Ed). ACSD.

Chad Hershock *(Carnegie Mellon University) is the Director of Evidence-Based Teaching & DEI Initiatives at the Eberly Center for Teaching Excellence & Educational Innovation. Chad leads the CTL's teams of teaching and assessment consultants.* **Jacqueline Stimson** *(Carnegie Mellon University) is a Senior Teaching Consultant at the Eberly Center for Teaching Excellence & Educational Innovation.* **Laura Ochs Pottmeyer** *(Carnegie Mellon University) is a Data Science Research Associate at the Eberly Center for Teaching Excellence & Educational Innovation.* **Michael C. Melville** *(Carnegie Mellon University) is a Senior Data Science Research Associate at the Eberly Center for Teaching Excellence & Educational Innovation.* **Jessica Harrell** *(Carnegie Mellon University) is a Senior Teaching Consultant at the Eberly Center for Teaching Excellence & Educational Innovation.* **Emily Daniels Weiss** *(Carnegie Mellon University) is Senior Teaching Consultant at the Eberly Center for Teaching Excellence & Educational Innovation.* **Marisella Rodriguez** *(University of California Berkeley) is a Senior Consultant at the Berkeley Center for Teaching Learning.* **Sophie le Blanc** *(Carnegie Mellon University) is a Senior Teaching Consultant at the Eberly Center for Teaching Excellence & Educational Innovation.* **Alexis Adams** *(Carnegie Mellon University) is a Senior Instructional Teaching & Learning Specialist at CMU-Africa.*

Appendix A

Workshop Title & T-SEDA approach	Learning Objective	T-SEDA, version A	T-SEDA, version B
Teaching Metacognitive Skills T-SEDA approach: vary cover story	Identify teaching strategies that teach students metacognitive skills. Explain why these teaching strategies are effective for teaching metacognitive skills.	Sasha is a junior in your upper-level engineering class. One of the main learning objectives of the course is for students to apply course concepts that they learned in lower-level classes to real world scenarios. Sasha is very distraught because they failed the first exam, and that's never happened to them before. When you ask how they studied for the exam, they say, "I studied the same way that I always have - I reread the textbook, underlined/highlighted everything, and looked at my lecture notes. I learn best through repetition, so reviewing the material is a useful study strategy for me." 1. What would you do as an instructor to promote Sasha's metacognitive skills? 2. Why would those strategies help?	Farrah is a first-year student in your first-year writing seminar. She approaches you after class to discuss her paper grade. Her paper was typical for first-year writing: it lacked a clearly articulated argument, and there was only weak evidence to support it. She explained that she was a "gifted" writer who had always received As on her high school English papers. She made it clear that there must be some mistake in this paper's grade because her mother, a high school English teacher, had read the paper over the weekend and thought it was very strong. Farrah admitted that she had started this assignment the night before it was due, but insisted that she worked best under pressure, saying, "That's just how my creative juices flow." 1. What would you do as an instructor to promote Farrah's metacognitive skills? 2. Why would those strategies help?
Creating a Supportive Classroom Climate from Day 1	Identify and implement strategies for creating an inclusive	For creating an inclusive space where students feel welcome:	For creating an inclusive space where students feel welcome:

T-SEDA approach: Vary questions	classroom environment that supports learning for all students. Identify strategies for getting to know your students.	1. Why do you think integrating a range of relevant examples could be an effective strategy? 2. Why do you think allowing students time to think and write could be an effective strategy?	1. Why do you think emphasizing a range of viewpoints could be an effective strategy? 2. Why do you think having students work in pairs could be an effective strategy?

Appendix A. Sample, counterbalanced T-SEDA assessment items from two CTL workshops that vary the cover story or questions for the same learning objective.

Appendix B

Evolution of T-SEDA Items Across Three Iterations of a Workshop on Working Well One-on-One During Office Hours

NB: The bolded phrase illustrates how we adapted a particular cover story over several iterations of the assessment instrument without changing the core component (i.e. this student demonstrates the Multiplicity stage of intellectual development). It was not bolded in any assessment instrument participants saw.

1. Fall 2018: Open Ended, Qualitative Samples of Students Responses

Assessment

The following quotation was recorded from an undergraduate student during office hours with a TA. Please read the quotation and then respond to the prompts below.

"I've noticed that the professor says a lot of things based on her own experience. I think that's good in a lot of ways, though, because it makes me think, **'Not everything has a right answer.'** If one possible solution made more sense to me personally, that's probably the way I would go."

Imagine you are working with this student one-on-one during office hours:

1. How does this student view knowledge?
2. How could you, as an instructor, best support this student's development as a learner, while working with him/her one-on-one?
3. What is your rationale for your answer to item 2 above, based on how students develop intellectually?

Examples of Participant Responses on Post-Assessments

Participant #19

1. "Multiplicity"
2. "Employ strategies for multiplicity → relativism"
3. "Basis for ideals is super important."

Participant #29

1. "For receiving truth and the answer."
2. "Depending on the problem give different answer."
3. *[left blank]*

Participant #43

1. "Multiplicity"
2. "Showing the weaknesses of one option and importance of taking being open to new ideas."
3. "Because I think the student stage of knowledge is multiplicity. I try to encourage him to move the next stage."

2. Spring 2019: Multiple Choice Question, Version 1

Assessment

After a recent lab report is graded and returned to students, a student says to the TA in office hours: "I've noticed that the professor often says a lot of things based on her own experience as a biochemist. I feel like when it comes to completing these labs, my way is just as valid as how she is telling us to do it. **After all, not everything has a right answer, right?**"

Which of the following is the MOST effective strategy for supporting this student's intellectual development while working one-on-one?

A. The student needs to recognize that the way we are doing things in class is important to master. I would explain to the student that the professor's years of experience have provided her with insights that cannot be overlooked. (27%*)
B. It sounds like this student is beginning to use her intuition more effectively and building an approach to the labs around other things she knows about science. I would encourage the student to articulate why she took that approach. (40%)
C. I agree with the student that there is more than one way to test a hypothesis and this is a critical element of the scientific method. I would explain how there will be chances to explore this in more advanced courses or through undergrad research opportunities. (20%)
D. I'm worried the student is not prioritizing the feedback from the instructor. I would review their graded lab report with them and

coach them to identify the skills needing the most attention for improvement on the next lab assignment. (13%)

*Values in parenthesis indicate the proportion (%) of respondents (n=15) selecting each distractor.

3. Fall 2019, Spring 2020: Multiple Choice Question, Version 2

Assessment

After a recent lab report is graded and returned to students, a student says to the TA in office hours: "I don't understand why I have been graded down on this report. I've noticed that the professor often says a lot of things based on her own experience as a biochemist. I usually don't take notes on those things or use them to prepare my labs. I feel like when it comes to completing these labs, my way is just as valid as how she is telling us to do it. **After all, not everything has a right answer, right?** "

1. From an intellectual development standpoint, what is the best description of what is going on?

A. The student is taking poor notes for the labs.
B. The student is effectively exploring the scientific method.
C. The student doesn't value the instructor's experience as a biochemist.
D. The student thinks that any approach is just as valid as the next.

2. Which of the following is the best strategy for supporting the student's intellectual development?

A. I would suggest the student accept the instructor's feedback due to their vast experience.
B. I would coach the student to articulate why they took that approach.
C. I would encourage the student to wait for more advanced courses to take that approach
D. I would point out specific strategies the student should practice before the next lab assignment.