CAPABILITIES: ADVANCING RESEARCH AND PRACTICE IN TEACHING AND LEARNING

Working Group Report #3

Working Group: **Kristinn Már Ársælsson** (principal author), Meifang Chen, Mengtian Chen, Kim Hunter Gordon, Pascal Grange, Kai Huang, Yitzhak Lewis, Junyi Li, Andrew MacDonald, Ben Van Overmeire, Bill Parsons, Noah Pickus, Renee Richer, Ira Soboleva, Mark Spaller, Daniel Weissglass, Jiaxin Wu, Ying Xiong, Xiaoqian Xu, Haiyan Zhou

Duke Kunshan University
Institute for Global Higher Education

CAPABILITIES:

ADVANCING THEORY AND PRACTICE

IN TEACHING AND LEARNING

Contents

INTRODUCTION

The Vice Chancellor for Academic Affairs (VCAA) charged an IGHE Working Group with examining the best theory and practices in teaching and learning that will enable Duke Kunshan University (DKU) to act as a leader in a) high-quality pedagogical experimentation and innovation and b) developing authentic assessment tools that can enable our decision-making about teaching and learning.

A central tenet of DKU's mission is to provide the highest quality undergraduate and graduate education. To that end, the university provides training to all incoming faculty, continuously offers feedback by experts, and equally weighs teaching excellence and research for tenure and promotion. Many faculty were attracted to the opportunity to participate in developing world-class teaching at this new university. A recent poll shows that, on average, faculty spend more than half of their working hours on teaching and related activities. Overall, there is good reason to believe that the quality of teaching is higher compared to many other universities.

Nevertheless, a survey of education science research and conversations among Working Group members suggest there are significant opportunities for improvement if DKU is to fulfill its highest aspirations Many faculty are eager to learn more about recent advances in education research and to apply those lessons in their classrooms. And despite strong evidence in favor of more effective pedagogies, some old habits endure. The Working Group reviewed this evidence and how faculty and students at DKU can benefit significantly from a systematic evaluation and further development of teaching and learning standards and support. Identifying what are core features of evidence-based high-quality teaching and learning is essential to design effective and efficient training programs, fostering a culture of excellence, and upholding standards.

The Working Group explored several well-established best practices, including active instead of passive learning, retrieval for long-term retention, self-regulated learning skills, various types of feedback, and teamwork. While most faculty are familiar with these approaches, Working Group members gained new insights about all of these practices. Additionally, recent research suggests several opportunities to further test and integrate innovative approaches to teaching. For example, while Intelligent Tutoring Systems (ITS) are not new, with the continuing development of Artificial Intelligence (AI) there is good reason to identify how ITS and human-directed teaching can work together. (A separate IGHE project will focus on the broader implications of AI for how we live, work, and learn.) Other promising approaches include experiential learning, authentic assessment, making errors a central focus, and throwing students in at the deep end.

In the following sections we describe what the Working Group learned and discussed, including topics such as learning theory, feedback, and intelligent tutoring systems. Next, we provide a brief overview of how these themes apply to DKUs context. For example, DKU has several uncommon features, including seven-week courses, a highly diverse campus, and interdisciplinary majors. Lastly, we provide thoughts about opportunities for improvement.

WHAT DID WE LEARN?

The Working Group covered three topics: 1) learning theory, 2) extensively researched teaching practices, and 3) innovative and less researched approaches. This section provides a brief overview of key insights from each section.

Learning Theory

Many pedagogical training programs center on *teaching,* understood as what the instructor does. However, any student-centered approach must give attention to *learning*. Course goals are *learning objectives,* not teaching objectives. Teaching, thus conceptualized, is to facilitate learning. The group reviewed several summaries from the study of learning theory. Research suggests several strategies that can help promote long-term retention. A few examples include retrieval, spaced out practice, dual coding, and interleafing (see box for details).

> A simplified summary of the learning theory literature suggests that using information repeatedly with limited cues, in different contexts, using both verbal and visual representations creates more numerous and durable brain pathways:[1,2]
>
> - Studies suggest that **retrieval and recall** from memory outperforms many other learning approaches, including rereading, note taking, and listening to lectures. This phenomenon is sometimes called "the testing effect", i.e. many studies show that the act of taking a test (which requires recall from memory) promotes long-term retention. Scholars recommend using formative assessment as a teaching method.
>
> - **Elaboration** requires students to provide multiple details when describing and explaining concepts, problems, frameworks etc. Adding details forces students to add depth to their understanding and think through application in different contexts.
>
> - Retrieval practice needs to be **spaced out**, i.e. repeated over a period to increase the number of connections and strengthen them. This finding has led many education scientists to argue that "less is more", i.e. courses that quickly jump from one topic or skill to another create fleeting memories that quickly fade.
>
> - Research suggests that providing **concrete examples** of abstract ideas can facilitate a better understanding and stronger long-term retention. Applying concepts, theories, and approaches to real-world problems can also spur interest as students realize the practical value of learning abstract ideas.
>
> - Using **dual coding**, i.e. using verbal and visual representations of a concept, can also stimulate long-term retention as information is stored in different formats (and places in the brain). Note though that there is strong evidence against *learning styles theory* – that different learning styles are more or less effective

for students. The evidence suggests we learn from using *both* verbal and visual channels of information, not that we learn more from one or the other.

- Finally, cycling through multiple subjects – what scholars call **interleafing** – supports long-term retention and transfer. Interleaving is often contrasted with "blocked practice" where students focus on one topic or skill at a time. By alternating between e.g. concepts, students learn to differentiate between them, which deepens understanding and improves retention.

Research on long-term retention is limited but one study showed that a year later, medical students retain about 60% of what they learned.[3] Some Working Group members felt that was surprisingly high while others pointed out that the average score on the repeated test was a failing grade, around 50%. Faculty members raised concerns that retention rates might be lower in many other disciplines given selection-effects and strong emphasis on memorization training in medical schools. As one participant said about their own field, "they forget everything during the summer". One idea that came up during our session was to have students, at the beginning of a course, recall what they learned in the previous course. That information could then be used by the instructor to adjust the level for the current course and review topics and add training for skills that are still developing.

Next, we explored self-regulated learning, including what scholars call self-regulated learning or "metacognition" (see box for details). Research shows that metacognition helps improve learning among students and that it is separate from intelligence. Unfortunately, self-regulated learning is unequally distributed among students and commonly at a low level. Classic examples include limited planning and "cramming" before an exam. Many WG participants felt that many students are too passive, disorganized, and ineffective in their self-regulated learning, e.g. at most read and reread while taking ineffective notes shortly before exams or assignments. Much of our discussion was thus centered on the importance of training metacognitive skills through assignments and reflection. Just as faculty can cram too much content into their courses, students often do not reflect on their own learning and actively build up their knowledge and skills. In other words, too much content and insufficient learning skills mean that universities, paradoxically, can often lack ideal contexts for learning.

Metacognition is commonly split into three parts: planning, monitoring, and evaluating:[4]

- Students should set themselves goals for each course and plan how they aim to achieve them. An important part of this process is making decisions on what kinds of learning approaches they'll employ to achieve said goals. Often students spend little time thinking about the effectiveness and efficiency of their approach to learning.
- Next, they monitor their progress towards the goals and consider whether adjustments need to be made. Can I find a more efficient approach? What support is available?
- And finally, after the course, students evaluate what they learned, how effective their learning was, and what they need to reach the next level. Another important

part of self-regulated learning is emotion-management, e.g. the ability to exert a sustained effort when learning is "hard" and "outside one's comfort zone".

The Working Group agreed that further attention to self-regulated learning has great potential for improving the quality of education at DKU. This would involve further development of how to facilitate asynchronous/out-of-class learning in classes at DKU. By default, students only spend about a quarter of their course time in class with the instructor. This presents an opportunity for DKU that has yet to be fully leveraged. The Working Group discussed strategies about how to equip students with metacognitive skills so that they can regulate and adjust their learning outside of the class time to achieve better outcomes.

Extensively Researched Teaching Practices that Foster Active Learning

Much research shows that active learning outperforms passive (e.g. lectures) for various learning objectives. Active learning commonly encompasses at least three aspects of learning: behavioral, cognitive, and social. The core idea emerged from learning theory, i.e. that long-term retention is more likely when students are actively engaged, e.g. when they apply, evaluate, create, deliberate, and collaborate. Analysis of 225 studies in science engineering and mathematics found that compared to passive instruction, student performance improved under active learning as well as reducing failing rates.[5] Similar findings have been reported for Asian students in STEM subjects,[6] humanities and social sciences,[7] and non-academic outcomes like confidence and interpersonal skills.[8]

However, sometimes and/or some students feel like they learn less from active learning than passive and prefer lectures[9] (although note that many studies report a positive attitude towards flipped classrooms/active learning). This feeling can lead to disengagement which might limit learning benefits. Scholars hypothesize that this mismatch between the feeling of and actual learning can be due to preconceived notions of what counts as learning *or* due to an increased effort – because active learning is "harder" students might think they are not good at it and mistake that feeling for lack of learning. Faculty should address such concerns directly and provide students with sufficient description and rationale when they employ active learning strategies (or any approach that is unconventional and/or requires sustained effort). It is worth noting that several studies show that active learning courses get higher course evaluations compared to lecture-based courses.[10]

<u>Use of Feedback</u>

The group spent most of their time discussing various types and **the use of "feedback"** in courses. Feedback in some form is nearly ubiquitous in higher education. Ranging from a single (summative) number or letter grade to detailed (formative) description of e.g. what went wrong, how to make it right and avoid similar mistakes, and what to do next (see box for details).

Research shows that feedback can facilitate learning but is sometimes not even reviewed by students. This might be because the feedback doesn't include sufficient details for students to understand what they did wrong, should have done, or what to do next and thus don't see the reason to review it, especially when there are no opportunities to improve.[11] Because giving feedback can be time intensive, faculty were interested in learning that peer-feedback can foster learning and sometimes provide similar effects to receiving teacher feedback. However, the

mechanism might be reversed, i.e. for students it's the *giving* feedback that promotes more learning than receiving it. Peer-feedback might also force critical thinking because they receive information from a source they don't blindly trust.

> Research shows that feedback can have a positive impact on learning but also reduce confidence and limit risk-taking. Key problems include limited or untimely feedback and/or students do not review it. Feedback has multiple functions. Overall, it is aimed at reducing a "discrepancy between current understanding and performance and a goal"[12]. Beyond that, it can also help improve self-efficacy, metacognition (see box above), and guide students to levels beyond the course LOs. Scholars have identified different **types of feedback**.
>
> Most feedback is arguably about what students did correct and what they got wrong. However, research suggests other types are equally important while less prominent. For example, students sometimes misunderstand what the goal is or struggle to figure out what the correct answer is. Another dimension is the process, i.e. how to arrive at the correct outcome. Feedback can help students identify why they made a mistake and how to prevent similar mistakes in the future. Students sometimes acknowledge that they got something wrong, but struggle to figure out how to improve in the future.
>
> Relatedly, students often appreciate guidance on how to reach the level beyond the course goals and identifying further learning opportunities. Thus, feedback can be directed at e.g. the task (e.g. assignment performance), process (how to build understanding and improve performance), and self-regulation/metacognition (e.g. planning, monitoring, emotion management).

Use of Group and Teamwork

Significant research shows that group- and teamwork can facilitate learning. Teamwork and interpersonal skills are also increasingly important at firms and institutions around the world. A 2015 survey of employers and university students found that 96% of employers and 94% of students agreed that "all college students should have educational experiences that teach them how to solve problems with people whose views are different from their own".[13] Teamwork is currently popular at both private firms and public institutions.

> Teamwork has been a feature of human cooperation since early human kin-based tribes. In fact, scholars argue some of the features (both good and bad) of modern teamwork have deep evolutionary roots.[14] Work groups (teams) have been studied for over a century[15] and significant research has been devoted to exploring group learning. Meta-analyses of studies show that group learning and teamwork can foster learning and on average outperform e.g. lecturing.[16] Group-based learning also helps train interpersonal skills and team management. Studies report "higher test scores" and "that students are [often] more highly engaged, [and better] prepared for class".[20] While scholars are still working through the details of effective group- and teamwork for learning (see more below), some studies suggest that students who have "interpersonal

difficulties" earn higher grades on group tasks despite reporting low satisfaction.[17]

Working Group participants were particularly interested in details about how to organize group learning in classes. Recent studies highlight some of factors that are likely to influence the effectiveness of team-based learning. For example, a meta-analysis found that tasks which include a group presentation at the end of a class have a smaller overall effect on learning compared to those that don't include the presentation.[20] Another example is that "people-centered" group leaders seem to be better at facilitating learning compared to "task-focused" leaders.[18] One possible explanation for both findings is that presentations and heavy focus on the outcome leads to an uneven distribution of work, i.e. the more assertive and/or better prepared students take the lead. As a result, only some students are active leading to less learning overall. This also fits research results which suggest that students learn more from group activities with greater "inter-dependence" among team members.

Innovation and Less Researched Approaches

To what extent are learning objectives hierarchical? While scholars have long theorized that learning objectives are less hierarchical than often assumed (e.g. Bloom's Taxonomy), only recently have scholars begun to empirically unpack the relationship. Recent research suggests that students can jump directly (or quickly) to higher levels of learning. The Working Group discussed a recent study showing a weak link between lower-order and higher-order learning. Preparing for a lower-order quiz didn't help much for a higher-order quiz and *vice versa*. However, mixed preparation boosted performance on both types of quizzes.[19]

That finding might suggest it's possible to learn lower-order and higher-order content and skills simultaneously or in tight succession. Many courses focus only on lower-order training and others on higher-order. In some cases, this might be ineffective because the lower-order learning might be lost (see long-term retention above) before getting to the higher-order learning and thus too much time spent on lower-order learning. On the other hand, assuming lower-order understanding in higher-level courses is a missed opportunity to improve retention and integrate the two levels.

Next, the group learned about so-called Intelligent Tutoring Systems (ITS). Perhaps the most famous modern ITS is Duolingo. These systems are "intelligent" because they monitor progress and move students through levels based on information about prior success. Just like a tutor would. Research shows ITS can outperform large class teaching and self-directed textbook work and even rival human tutoring (often considered the most effective pedagogical approach).[20] While ITS systems have become popular outside of academia and are commonly used in e.g. teaching coding, they are relatively rarely used in academia. Some Working Group participants employ ITS or computer-assisted learning systems in their own courses, using e.g. Coursera. However, faculty are not confident of their effectiveness. Perhaps more important, it can be time intensive to find and curate ITS' because they were designed for a different curriculum. Nevertheless, participants saw great potential in using ITS. They also aired concerns that in the future they themselves might be replaced by ITS.

Intelligent Tutoring Systems are computer software that models learners' performance to provide individualized tutoring.[21] ITSs monitor the progress of

> the learner and propose exercises depending on prior progress and mistakes. These systems often use multiple types of feedback (see above) to "provoke cognitive, motivational or metacognitive change".[24] Some also use "gamification" to encourage further engagement and development. They've been employed in various fields (language, math, programming, law etc.) and at education levels. A 2014 meta-analysis found that ITSs was "associated with greater achievement in comparison with teacher-led large-group instruction (g = .44)".[24]

There is growing interest in what scholars call "authentic assessment", i.e. assessing students' ability to apply what they've learned to common tasks or problems they'll face in life or at work. Authentic assessment is aimed at reducing teaching, learning, and assessment that is limited to content or skills that are constrained to the classroom. Thus, it promotes "transfer" – the ability to use what you learn at school outside the classroom, e.g. at work or in life – and constrains what is commonly called "schooling". Critiques of schooling date back centuries, for example, Michel de Montaigne criticized education in the 16th century for training students not for life, but school (*non vitae sed scholae discimus*).

Broadly, scholars have long called for more research on assessment. One concern is that much assessment is not only inauthentic but also lacks validity – does not capture what it is intended to measure – and reliability – results are not consistent across and within individuals. Any experimentation with authentic assessment should also take those concerns into account.

> **Authentic assessment** "calls for students to utilize the same set of knowledge, competencies, and attitudes that they should apply […] in real-life".[22] Instead of only testing knowledge and retention, authentic assessment measures (and trains) practical application. Ideally, it replicates what students need to do in professional and/or personal settings. For example, faculty have used role-play to train and assess medical students' communication skills in simulated settings with a patient.[23] One nursing student commented that the authentic "assessment made me review lectures and textbooks more than I would normally". Another example is a study of authentic assessment for students in a management program. The students were asked to prepare a report based on a real-world challenge. The study found that the assessment not only equipped students "with leadership knowledge and skills, [but] more importantly, [helped] develop their understanding as leaders".[24]

While scholars and practitioners alike have long acknowledged that **learning from mistakes can be powerful,** systematic research of learning from errors is relatively underdeveloped. Recent studies suggest that focusing on errors can be comparably *effective* but more *efficient* than traditional approaches where errors are noted and/or the correct answer/approach provided by the instructor.[25] Relatedly, research suggests that making "deliberate errors" – what scholars call the "derring effect" – can be a particularly effective learning approach.[30] More work is needed to confirm the effectiveness and efficiency of these approaches, but scholars suggest that learning from mistakes can improve metacognition and reduce the probability of repeated errors. Similarly, the derring effect forces students to know and understand the material/skill at

a good level *and* creates additional brain pathways – what something is not – which increase probability of later recall.

> **The "derring effect"** is when students are asked to commit deliberate errors and then correct them. Research suggests that derring can outperform learning by generating correct answers for retention and transfer – applying the content or skills learned to other contexts.[26] Further study of why derring seems to be effective because students a) need to confidently know the correct answer, b) identify similar but incorrect answers (contrast), and c) reflect on *why* incorrect answers are unjustified.

DUKE KUNSHAN UNIVERSITY CONTEXT

This section summarizes some of the challenges and opportunities faculty discussed in relation to universities in general and DKU more specifically. Participants noted that DKU faculty might have more experience and advantage over other many other universities in several areas. One is "modularity", i.e. that most classes at DKU are taught in a seven-week schedule. Another is "experiential learning" which has its own space in DKU's curriculum. A third is small but relatively diverse classes. Finally, DKU offers an interdisciplinary curriculum.

Modularity. Almost all courses at DKU are taught in a seven-week structure whereas 14 weeks are more common. Thus, there is valuable experience and opportunity to explore the benefits and problems of modularity, i.e. the effects of time spent on a course. Faculty raised several concerns and challenges from their experience of the seven-week format. Some argued that they are not able to cover the necessary content in seven weeks and the shorter format induced anxiety among students. Against that, some argued that the shorter format is more effective owing to its intensity. Additionally, it might better train students to develop stronger learning habits. A few faculty noted that some DKU students say that Duke courses are "easy" compared to DKU courses.

Experiential Learning. DKU's curriculum values experiential learning and has built in reserved space for it on Fridays. Thus, DKU faculty are encouraged to incorporate experiential learning into their course design. Like authentic assessment, experiential learning aims to provide training in real-world contexts and link understanding to pressing issues. Research suggests that experiential learning can boost learning (see below). Some faculty at DKU employ experiential learning activities while others do not. Participants agreed that there is much promise in using experiential learning but also that faculty face challenges and need more training or support to implement it in their courses.

> Broadly, experiential learning is learning by experience. It can take many forms, inside and outside the classroom, and often includes self-reflection on said experience. While experiential learning has a long history and has been used widely, relatively few robust studies have tested its effectiveness. A 2019 meta-analysis identified over 13,000 articles about experiential learning but only 89 studies with a control and treatment group. Analysis of these studies suggest that experiential learning can improve learning outcomes (d = .43) compared to traditional learning environments. The largest gains were for cognitive development, compared to personal and social outcomes.[27] Future research

needs to identify the impact of experiential learning compared to active learning. How much is the "experience" part adding to the fact that most experiential learning is active learning? What kinds of experiences are most effective and efficient?

Small and Diverse Classes. Most classes at DKU are relatively small, have fewer than 18 students, many fewer than 10, and only a handful more than 30 students per instructor. Currently DKU does not have any major courses with more than 100 students per instructor, which is common in introductory courses at many universities. Faculty are thus able to use and experiment with teaching and learning activities that allow for significant interaction between the instructor and a large proportion, if not all, of the students. In large classes, faculty are limited in how much feedback they can provide, and the time given to interacting with individual students. Furthermore, DKU's student body is more "international" compared to most universities. This offers unique opportunities to engage students in dialogue about cross-cultural perspectives, interests, and problems. It can also be challenging because students' prior education experiences and expectations as well as their skills vary significantly.

> **Class Sizes and Diversity**. A large majority of DKU classes are capped at 18 students (per faculty member). However, many courses have <10 students while a few have more than 18 students. Currently the student to full-time faculty ratio at DKU is about 10:1. According to the National Center for Education Statistics, the average undergraduate class size in the U.S. is 25-30 students. However, it is common that introductory classes are significantly larger, some with several hundred students. The distribution of students per faculty member also varies significantly across majors and disciplines. Additionally, the proportion of international students to native Chinese students at DKU is about 2:1 whereas in the U.S. the proportion is about 20:1. Thus, DKU is proportionally much more diverse in terms of nationality than most U.S. universities. In terms of race, however, DKU is more like U.S. universities, with more than half of the students being Asian or non-Hispanic White respectively.

Interdisciplinary Curriculum and Teaching. DKU does not offer traditional majors that neatly align with disciplines. Nor does it have traditional structures for organizing faculty. Some courses are part of different majors and many faculty teach across disciplines. Thus, there are opportunities to experiment and research interdisciplinary teaching and learning. However, there are also several challenges. One is the legibility of DKU's majors to other universities and workplaces. Another is whether and then how to guarantee necessary (sometimes disciplinary) training that is commonly organized through a sequence of courses within a discipline. Third, many faculty were not trained in interdisciplinary teaching, nor do they participate in interdisciplinary research. Thus, they find it difficult to design interdisciplinary courses, teach outside their discipline, and/or engage in interdisciplinary research. It was not clear to Working Group participants how to achieve DKU's vision of training the next generation of scientists and leaders with an interdisciplinary training and perspective. A separate IGHE project has been launched to focus on this topic.

Challenges. The Working Group participants also raised several emerging challenges to teaching at DKU. First, the growth of generative AI in recent years has caused concern among

faculty. One issue is plagiarism and students using AI software to complete or revise assignments. Another is what kinds of skills and habits AI will replace. Is there course content or training that is already or will soon become irrelevant in a few years? For example, how important is it to teach coding from memory and by hand, as opposed to being able to use AI assisted coding? Second, participants raised concerns about how time intensive high-quality teaching can be – especially with extensive Signature Work and advising. Relatedly, some faculty noted the added complexity of simultaneously delivering world-class teaching and leading research that engages real-world problems at the same time. Third, while the Working Group faculty were enthusiastic about learning more from education science, some were apprehensive and even daunted by the task of navigating an already large and growing research literature. There are many viable approaches, making it difficult for faculty to identify which one is the most effective and efficient for the LOs and student composition.

Finally, the Working Group spent significant time discussing whether and then how to use lecturing as a mode of instruction. As noted above, a large body of research shows that active learning outperforms passive learning. In education research, lecturing is commonly defined as a passive teaching approach. However, participants raised questions like what is and what isn't a lecture? What kind of lecturing might be useful and appropriate? Are there strict time limits on lecturing? Under which conditions is lecturing necessary and justified? For example, one faculty noted that in some cases – particularly for interdisciplinary courses – there is a lack of available textbooks and other course materials. Under these conditions, lecturing might be necessary (at least as supplementary materials). Others warned that we should not equate every instance where the instructor speaks with lecturing. For example, when faculty review student assignments and outcomes from activities, they are providing actionable feedback which is more akin to tutoring (which is known to be a very effective approach). Nevertheless, it is important to remember that long lectures where students are passive recipients of information are on average less effective compared to a multitude of available active learning and teaching approaches.

FUTURE DIRECTIONS

Building on what the Working Group learned and discussed, this section identifies opportunities to strengthen and build on DKU's mission of providing world-leading training for the next generation of leaders and scientists. First, there is an opportunity to more broadly integrate well-established best practices of teaching and learning. To that end, identifying what comprises a set of pedagogical best practices is crucial. At the same time, expanding systematic experimentation with emerging technologies and pedagogies. Next, valid, reliable, authentic, and comparable measures of learning objectives can help facilitate integration and evaluate experimentation. Finally, more attention to efficiency in relation to effectiveness – how to maximize learning given time and resources can maximize impact while addressing time constraints.

World-Class Teaching and Learning:

On average, university faculty have little formal training in teaching and learning theory. The Working Group discussions identified several well-established practices that there is reason to believe are unequally distributed across courses at DKU. These include e.g. long-term retention

strategies, metacognition training and structured asynchronous learning, group activities, and intentional use of various types of feedback. Even faculty who actively engage with the Center for Teaching and Learning find it difficult to navigate the extensive education science literature.

- The Working Group participants called for an accessible summary of well-established pedagogical practices. Defining the key components of world-class teaching can also guide the development and implementation of e.g. peer-evaluation and faculty professionalization.
- Relatedly, there are various strategies available to integrate best practices at scale. One is to include them in peer-evaluation rubric which provides incentive for faculty to seek out training and incorporate them into their courses. Thus, peer-evaluations would reflect best practices identified by a significant body of research.
- Another is to create an online diagnostic tool to identify which kinds of synchronous and asynchronous activities and assessment are effective for specific LOs, content, and skill training. Such a tool could provide more detailed information that is curated according to the specific context of each course.

Finally, having a comprehensive overview of what are considered best practices of world-class teaching and learning habits makes it is easier to develop CTL programs and student training. Such an overview could form the basis of the one-year colloquium for incoming faculty. Such programs can be in-person or online self-paced modules. For example, the Center for Teaching and Learning is currently collaborating with faculty to develop a self-paced online module to train metacognition among students.

Exploring Emerging Technologies and Pedagogies:

Currently, DKU faculty are involved in various exploration and experiments around teaching and learning. CTL for example, offers funds and support for faculty education research. However, there are no systematic research programs on approaches that either fit DKU's mission or extant research suggests being impactful (either increase effectiveness or efficiency of teaching and/or learning).

Several themes that emerged through the Working Group might be worthwhile for further research and experimentation across courses at DKU. These include mixing lower- and higher-order training, Intelligent Tutoring Systems (ITS) and generative artificial intelligence, authentic assessment, and teaching from errors. We discuss each in turn below.

In many disciplines, content and skills are taught in an organized sequence. For example, basic concepts, key theories, and foundations of methods come before evaluation, design, and application of research. Another way to characterize the sequence is that lower-level courses teach "understanding" and higher-level courses "application". However, some research suggests that there is a weaker link between lower-order and higher-order learning than often assumed. Another piece of evidence is that much university learning happens in the first two years. The third suggestive evidence is that long-term retention is difficult to attain (if we don't use our skills or recall facts on a regular basis, they quickly fade). Thus, some sequences that are built into university curriculums might be ineffective – students forget basic training before

getting to higher-level application – or inefficient – students can gain understanding and learn basics while learning higher-order skills like application and evaluation.

Research shows that ITS can rival human tutoring, i.e. provide high impact teaching and learning. However, these systems are rarely used in higher education. Faculty who use them are often not trained in how to select or evaluate such systems. Nevertheless, ITS has the promise to provide training at various levels at scale, leaving more time for faculty to provide feedback and training that requires their expertise and/or physical presence. Similarly, the recent development of generative AI suggests multiple ways to improve both teaching and learning. AI might be used for rote tasks like grading, again saving time for faculty to stay in the zone of proximal development. A systematic review of current use of ITS and opportunities might lay the groundwork for integration into courses at scale.

DKU's mission is to enable students to "lead purposeful and productive lives" and prepare them for "professional, intellectual and societal leadership roles across the globe". Unfortunately, much higher education assessment does not directly measure or aid in training skills that students need or use after graduation. What scholars sometimes call "schooling". Against that, many scholars argue for further use of "authentic assessment" – where students are assessed on their ability to apply what they learn to real-world tasks.

Lastly, recent research suggests that learning from errors and derring (deliberative errors) can be effective. Studies suggest this approach can help with long-term retention and is more efficient than traditional approaches that focus on getting everything right in the first place. There is also reason to suspect a focus on errors can help students more easily step outside their comfort zone, improve self-efficacy, and even strengthen critical thinking skills.

Assessment of Outcomes

There are four reasons to improve, expand, and develop systematic assessment of learning outcomes.

First, some research suggests that students don't learn much during their college years and that much (if not all) of that learning happens in the first year or two.[28,29] These findings suggest that at least some teaching and learning is limited; the measurements used in the studies are not valid or reliable; and/or other learning objectives are not captured. Given the quality of many Signature Work Projects at DKU it is hard to believe students don't learn much if at all and that it only happens in the first year or two. However, it is possible that much of the learning happens through the SW itself, i.e. not from coursework. In any case, valid and reliable assessment is needed to be able to evaluate how effective and efficient teaching and learning is within and across courses.

Second, many of the measures currently employed at universities are not valid or reliable. For example, a meta-analysis found **no** relationship between student learning and course evaluations.[30] Another concern is that most university faculty are not trained in designing LO measurements. As a result, some course assessments are not valid or reliable.

Third, some of the current assessments, even if valid and reliable, are not authentic, e.g. only measures whether students are able to recall information in a classroom setting. Any university

that wants to deliver on its claims to provide high-quality training for professional vocations would benefit from having measures of student improvement and performance on real-world tasks. In short, to be able to share best practice or evaluate program implementation we need valid, reliable, and comparable (if not standardized) assessment.

Fourth, arguably, most education science studies focus on effectiveness and do not explicitly test efficiency. Teaching strategies and learning habits are compared and measured across the same amount of time. This analytical strategy tells us whether there is an overall average difference, but not whether the impact of one approach compared to the other was achieved in the same amount of time. However, if we can reach a similar outcome in significantly less time, that approach might be preferred.

RECOMMENDATIONS

Building on what the Working Group learned and discussed, this section provides a few recommendations for DKU.

- Expand the integration of well-established best practices in teaching into faculty training and evaluation including creating an accessible summary of best practices to guide the development and implementation of peer-evaluation and faculty professionalization.
- Undertake systematic experimentation with emerging technologies and pedagogies including moving from lower- to higher-order learning, AI and Intelligent Tutoring Systems, and learning from errors.
- Develop valid, reliable, authentic, and comparable assessment of learning objectives.
- Focus attention on efficiency in relation to effectiveness – how to maximize learning given time and resources in a modular curricular structure.

REFERENCES

[1] Halpern, Diane F. & Milton D. Hakel. (2003). "Applying the Science of Learning to the University and Beyond: Teaching for Long-Term Retention and Transfer." *Change: The Magazine of Higher Learning*.

[2] Sawyer, R. Keith. (2022). An Introduction to the Learning Sciences. In R. Keith Sawyer (Ed.) *The Cambridge Handbook of the Learning Sciences*. Cambridge University Press

[3] Schneid, Stephen D., Hal Pashler & Chris Armour. (2019). "How much basic science content do second-year medical students remember from their first year?" *Medical Teacher,* 41:2, 231-233, DOI: 10.1080/0142159X.2018.1426845.

[4] MIT Teaching + Learning Lab, Metacognition. Retrieved from the MIT Teaching and Learning Lab on April 12th; Ohtani, Kazuhiro & Tetsuya Husasaka. (2018). "Beyond intelligence: a meta-analytic review of the relationship among metacognition, intelligence, and academic performance." *Metacognition Learning* 13:179-212; Perry, John, David Lundie & Gill Golder. (2019). "Metacognition in schools: what does the literature suggest about the effectiveness of teaching metacognition in schools?" *Educational Review 71*(4): 483-500.

[5] Freeman, Scott, Sarah L. Eddy, Miles McDonough, Michelle K. Smith, Nnadozie Okoroafor, Hannah Jordt, and Mary Pat Wenderoth. (2014). Active learning increases student performance in science, engineering, and mathematics. *PNAS 111*(23):8410-8415.

[6] Ting, Fridolin S. T., Ronnie H. Shroff, Wai Hung Lam, Raycelle C. C. Garcia, Chi Lok Chan, Wing Ki Tsang, Ndudi O. Ezeamuzie. (2023). A Meta-analysis of Studies on the Effects of Active Learning on Asian Students' Performance in Science, Technology, Engineering and Mathematics (STEM) Subjects. *Asia-Pacific Education Research 32*:379-400.

[7] Kozanitis, A., Nenciovici, L. (2023). Effect of active learning versus traditional lecturing on the learning achievement of college students in humanities and social sciences: a meta-analysis. *Higher Education* 86:1377–1394. DOI: 10.1007/s10734-022-00977-8.

[8] Bredow, Carrie A., Patricia V. Roehling, Alexandra J. Knorp, and Andrea M. Sweet. (2021). To Flip or Not to Flip? A Meta-Analysis of the Efficacy of Flipped Learning in Higher Education. *Review of Educational Research 90*(6):878-918. DOI: 10.3102/00346543211019122.

[9] Deslauriers Louis, Logan S. McCarty, Kelly Miller, Kristina Callaghan, and Greg Kestin. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *PNAS 116*(39):19251-19257.

[10] Tutal, Özgür., & Taha Yazar. (2021). Flipped classroom improves academic achievement, learning retention and attitude towards course: A meta-analysis. *Asia Pacific Education Review, 22*(4), 655–673. DOI: 10.1007/s12564-021-09706-9.

[11] Wisniewsk, Benedikt, Klaus Zierer and John Hattie. (2020). The Power of Feedback Revisited: A Meta-Analysis of Educational Feedback Research. *Front. Psychol*. 10:3087. DOI: 10.3389/fpsyg.2019.03087.

[12] Hattie, John & Helen Timperley. (2007). The Power of Feedback. *Review of Educational Research 77*(1):81-112. DOI: 10.3102/003465430298487.

[13] Hart Research Associates. (2015). Falling Short? College Learning and Career Success Selected Findings from Online Surveys of Employers and College Students Conducted on Behalf of the Association of American Colleges & Universities. Hart Research Associates.

[14] Yong, J., Park, G., & Spitzmuller, M. (2021). From the Savannah to the Corporate Office: The Evolution of Teams. *Small Group Research 52*(1), 33-67. DOI: 10.1177/1046496420960516.

[15] Mathieu, J. E., Hollenbeck, J. R., van Knippenberg, D., & Ilgen, D. R. (2017). A century of work teams in the *Journal of Applied Psychology*. *Journal of Applied Psychology102*(3), 452–467. DOI: 10.1037/apl0000128.

[16] Liu, S.-N. C., & Beaujean, A. A. (2017). The effectiveness of team-based learning on academic outcomes: A meta-analysis. *Scholarship of Teaching and Learning in Psychology 3*(1), 1–14. DOI: 10.1037/stl0000075; Swanson, Elizabeth., Lisa V. McCulley, David J. Osman, Nancy Scammacca Lewis, & Micheal Solis. (2019). The effect of team-based learning on content knowledge: A meta-analysis. *Active Learning in Higher Education 20*(1):39-50. DOI: 10.1177/1469787417731201; Tomcho, Thomas J. & Rob Foels. (2012). Meta-Analysis of Group Learning Activities: Empirically Based Teaching Recommendations. *Teaching of Psychology 39*(3):159-169. DOI: 10.1177/0098628312450414.

[17] Lavy, Shiri. (2017). Who benefits from group work in higher education? An attachment theory perspective. *Higher Education 73*:175–187. DOI: 10.1007/s10734-016-0006-z.

[18] Koeslag-Kreunen, Mieke., Piet Van den Bossche, Michael Hoven, Marcel Van der Klink, & Wim Gijselaers. (2018). When Leadership Powers Team Learning: A Meta-Analysis. *Small Group Research 49*(4):475-513. DOI: 10.1177/1046496418764824.

[19] Agarwal, P. K. (2019). Retrieval practice & Bloom's taxonomy: Do students need fact knowledge before higher order learning? *Journal of Educational Psychology 111*(2):189–209. DOI: 10.1037/edu0000282.

[20] Ma, Wenting, Olusola O. Adesope, John C. Nesbit, & Qing Liu. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology 106*(4):901–918. DOI: 10.1037/a0037123.

[21] Tlili, Ahmed, Soheil Salha, Huanhuan Wang, & Ronghuai Huang. (2023). Intelligent Tutoring Systems Examined in Social Experiments—Is the Magic Gone? A Meta-Analysis. *2023 IEEE International Conference on Advanced Learning Technologies (ICALT)*, Orem, UT, USA: 50-54. DOI: 10.1109/ICALT58122.2023.00020.

[22] Sokhanvar, Zahra, Keyvan Salehi, & Fatemah Sokhanvar. (2021). Advantages of authentic assessment for improving the learning experience and employability skills of higher education students: A systematic literature review. *Studies in Educational Evaluation 70*:1-10. DOI: 10.1016/j.stueduc.2021.101030.

[23] Craft, Judy, & Louse Ainscough. (2014). Development of an electronic role-play assessment initiative in bioscience for nursing students. *Innovations in Education and Teaching International 52*(2):172–184. DOI: 10.1080/14703297.2014.931241.

[24] Wiewiora, Anna, & Anetta Kowalkiewicz. (2018). The role of authentic assessment in developing authentic leadership identity and competencies. *Assessment & Evaluation in Higher Education 44*(3):415–430. DOI: 10.1080/02602938.2018.1516730.

[25] Metcalfe, Janet, Judy Xu, Matti Vuorre, Robert Siegler, Dylan Wiliam, & Robert A. Bjork. (2024). Learning from errors versus explicit instruction in preparation for a test that counts. *British Journal of Educational Psychology* 00:1–15. DOI: 10.1111/bjep.12651.

[26] Wong, Sarah S.H. (2023). Deliberate Erring Improves Far Transfer of Learning More Than Errorless Elaboration and Spotting and Correcting Others' Errors. *Educational Psychology Review 35*:16. DOI: 10.1007/s10648-023-09739-z.

[27] Burch, Gerald F., Robert Giambatista, John H. Batchelor, Jana J. Burch, J. Duane Hoover, & Nathan A. Heller. (2019), A Meta-Analysis of the Relationship Between Experiential Learning and Learning Outcomes. *Decision Sciences Journal of Innovative Education 17*:239-273. DOI: 10.1111/dsji.12188.

[28] Flowers, Lamont, Steven J. Osterlind, Ernest T. Pascarella, & Christopher T. Pierson. (2001). How Much Do Students Learn in College? Cross-Sectional Estimates Using the College BASE. *The Journal of Higher Education 72*(5):565–583. DOI: 10.2307/2672881.

[29] Liu, Ou L., Katrina C. Roohr, Jacob M. Seybert, & and Daniel Fishtein. (2022). Are Fourth-Year College Students Better Critical Thinkers than Their First-Year Peers? Not So Much, and College Major and Ethnicity Matter. *Educational Measurement: Issues and Practice 41*:64-79. DOI: 10.1111/emip.12430.

[30] Uttl, Bob, Carmela A. White, & Daniela Wong Gonzalez. (2017). Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related. *Studies in Educational Evaluation 54*:22-42. DOI: 10.1016/j.stueduc.2016.08.007.