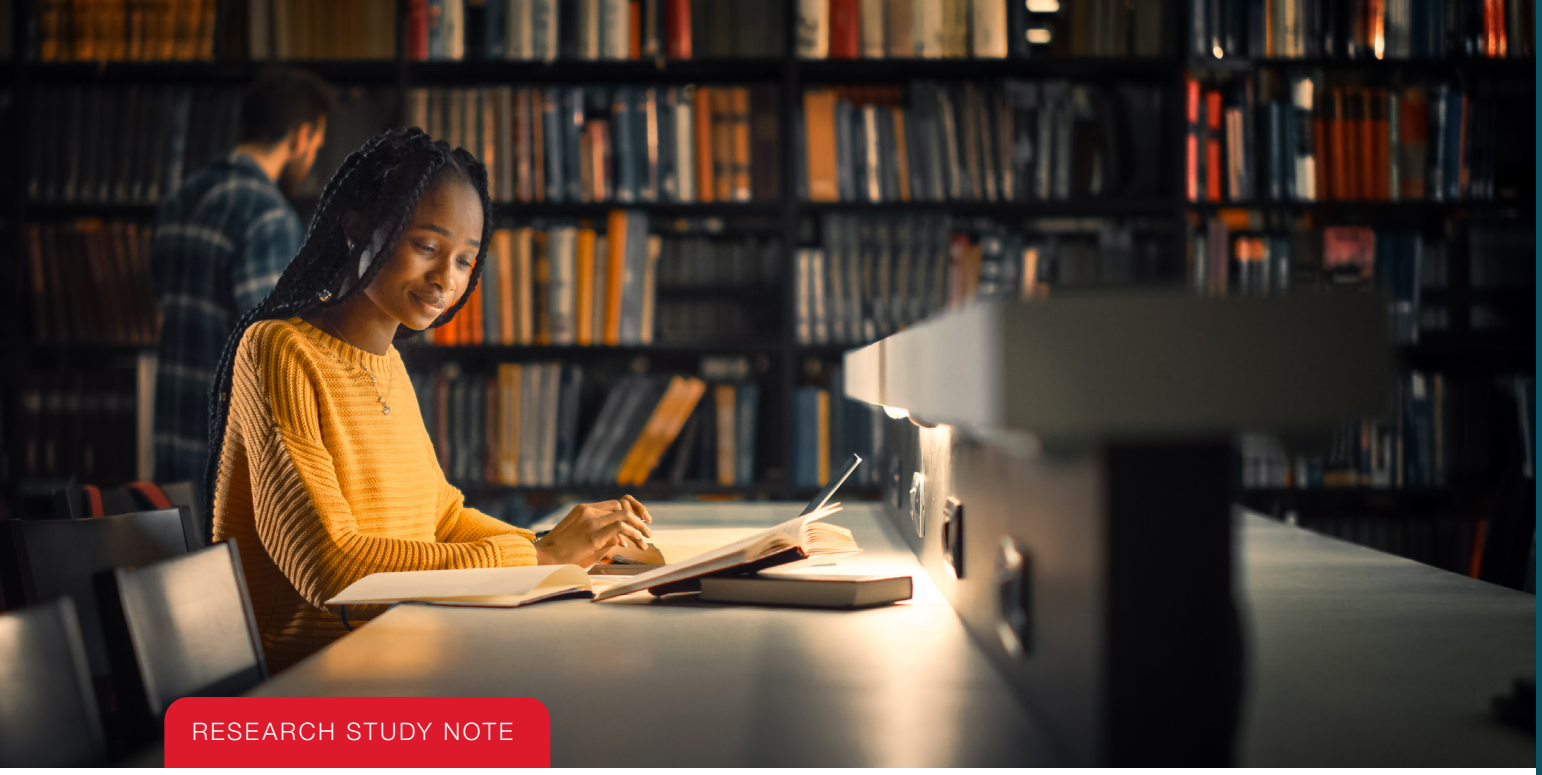


RESEARCH STUDY NOTE

Homework with Feedback

Key Findings:

- Course performance increased as students completed more Homework with feedback activities.
- Students who were the first in their family to attend college and students eligible for financial aid especially benefited academically from completing more Homework with feedback activities.
- Giving students the opportunity to practice their learning in a scaffolded environment may help close previously seen achievement gaps.

macmillan learning

BACKGROUND

A major goal for higher education is for students to gain the knowledge and skills to become lifelong learners in order to be successful in college and beyond. In order to reach this goal, students must have an understanding of what they know and what they don't know. This understanding is formed in large part by practicing particular knowledge and skills, receiving scaffolded feedback, and adapting when needed. Feedback can be given in a variety of forms including through instructors, classmates, digital tools, or the learner themselves through self-evaluation.

Previous research has found students who engage in more frequent assessment with quality feedback also have higher motivation, academic engagement, metacognitive skills, adaptive strategy use, academic performance, and likelihood of transferring their knowledge to new contexts (Cogliano et al., 2020; Han & Finkelstein, 2013; Shute, 2008). This type of assessment, frequently referred to as formative assessment, is a specific type of "assessment for learning" - it's not meant to be summative "proof" of learning, but rather support the learning process. Feedback is a critical component for formative assessment to support learning - it helps learners identify gaps in their knowledge/skill sets (Angelo & Cross, 1993; Bennett, 2010; Black & William, 1998). It also provides students with a path for continuous improvement - taking feedback, applying it, and continuing the path towards lifelong learning.

Feedback that can be personalized to individual students is especially important to ensuring an equitable learning environment. Furthermore, mixed assessment methods, which aim to reduce the impact of high-stakes exams and emphasize the significance of lower-stakes assignments, has been recognized as a means to create a more equitable learning environment and overcome obstacles for underrepresented students (Cotner & Ballen, 2017; Malespina & Singh, 2022).

The mechanisms that underlie this connection are still under investigation. One potential explanation is that active-learning pedagogy, which incorporates frequent low-risk assessments, establishes a structured environment where students can engage in regular problem-solving and practice, after instruction, while receiving feedback. Prior research has revealed that this pedagogy is linked to enhanced academic performance in all students, with particularly pronounced benefits for underrepresented students (Haak et al., 2011). This increased practice may boost students' self-efficacy, resulting in improved academic performance, which disproportionately benefits underrepresented students (Ballen et al., 2018).

Another potential explanation is that high-stakes assessments might be more prone to eliciting stereotype threat. Stereotype threat is the phenomenon where individuals are reminded of negative stereotypes that pertain to them, which can instill a sense of doubt in their abilities (Schmader, 2010). The heightened pressure to "prove" capabilities that accompanies high-stakes assessments is more likely to activate this mode of thinking. Consequently, individuals may expend cognitive effort beyond the immediate task at hand, ultimately leading to reduced

performance. Three groups of students commonly identified as experiencing stereotype threat in higher education encompass racial/ethnic minority students, first-generation college attendees, and students from lower-income backgrounds (Dennehy et al., 2018; John-Henderson et al., 2014).

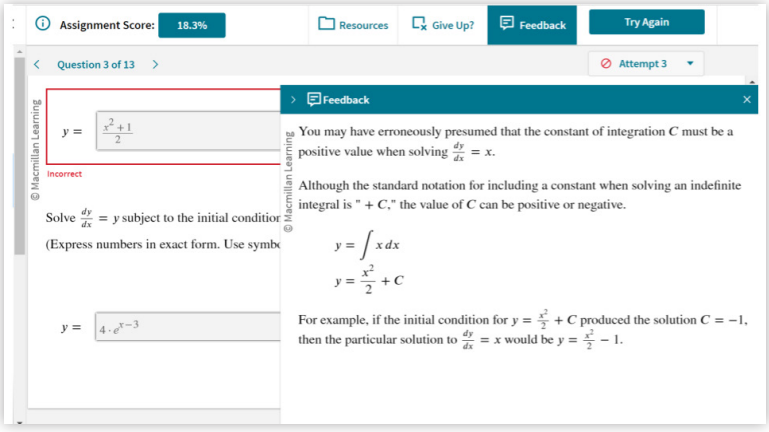
Macmillan Learning created an assessment tool that includes Homework with feedback (from here forward referred to as HwF) with the goal of providing students formative assessment with quality feedback. In order to examine its' impact on students' learning, Macmillan Learning funded a series of research studies, across six semesters (2019-2022) and 161 institutions, to examine the impact of HwF. Participating instructors were given training and implementation recommendations, but use of HwF activities was not required to participate in the study and implementation choices varied by instructor.

PRODUCT OVERVIEW

At Macmillan Learning, our vision is that all assessments are built intentionally, based on a learning framework, with the learner in mind. Our assessment tool, including HwF, was created based on a wealth of learning science research around three best practice principles: retrieval practice, spacing, and formative assessment with feedback. The act of retrieving information from memory through testing (i.e., retrieval practice) has been shown to lead to better memory of the material than rereading course text or notes (Roediger & Karpicke, 2006). Not only is it important to study and learn by testing yourself, it's also important to space out study sessions/testing. Spreading out study sessions over weeks rather than days has been found to relate to better performance on exams as well as long-term memory retention (Delaney et al., 2010). Formative assessment with feedback was discussed in the background section above.

Achieve Homework offers targeted feedback to help learners address misconceptions in real time. Students are able to access e-book resources as necessary, and review detailed solutions either when they reach the correct answer or if they give up. They are also able to receive a hint before submitting an item answer and receive targeted feedback to an incorrect answer, as shown in the example below. If a given student answer doesn't align with a typical misconception or incorrect answer, a default feedback response is given.

Example HwF Item



STUDY DESIGN

Ethics and Data Privacy

Prior to data collection, this study and the associated consent forms and instruments were reviewed and approved (found exempt) by the Human Resources Research Organization (HumRRO). HumRRO is an accredited, third-party Institutional Review Board organization with no affiliation with Macmillan Learning. Macmillan Learning seeks third-party review to eliminate any bias in the decision of the exemption. The data in this study, which are provided by the instructor and consenting students, are initially identifiable. However, once a random identifier is generated identifiable data are destroyed. Data is stored in secure storage locations, and access is permitted only to the primary investigator in the study.

Sample

The full study sample included participants spanning across six semesters from fall 2019 through spring 2022. This robust sample included 177 unique instructors teaching 333 courses. Eight different subject areas (chemistry, biochemistry, biology, calculus, precalculus, psychology, economics, and English) are represented in the study. Instructors came from 161 institutions across the United States and Canada. The sample included a range in institution and course sizes as well as course formats (i.e., face-to-face, virtual synchronous, virtual asynchronous).

The variation in participating institutions and instructors enabled a diverse student sample. The full student sample included 40% non-White or Asian, 24% who were first in their families to go to college, 66% who were eligible for financial aid, and 38% who had a high school GPA lower than 3.5 across a total of 9,803 participating students.

Methods

After consenting to be part of the larger research study, participating instructors were given brief training on HwF, practical information on where to locate and assign HwF to students within the Achieve program and its functionality. The HwF training was included as part of a broader training on Achieve given by a curriculum specialist and lasting approximately 45 minutes. If instructors requested follow-up training on any feature within Achieve, additional training was given. Participating instructors were not required to assign HwF to their students, enabling a more naturalistic implementation. Use of HwF was, however, observed and documented by the research team.

Students who consented to participate in the study granted researchers access to their course performance data as well as their HwF usage data. Furthermore, as part of participating in the study, students were asked to complete two additional surveys (beginning and end of semester) to share sociodemographic information, as well as general perceptions of Achieve.

DATA ANALYSIS

Statistical modeling (i.e., linear mixed model) was used to isolate the unique impact of completing HwF on student course performance (overall course grade and exam average grade as percentages). In order to partial out the unique impact of HwF, several factors were included in the model to control for other variables researchers thought would likely impact academic performance. The variables were:

- subject,
- course mean grade,
- course mean exam average,
- student college readiness (i.e., high school GPA, SAT/ACT scores),
- student gender,
- student race/ethnicity,
- first generation college student status,
- financial aid eligibility.

Including these variables in the model was an attempt to equate students on background variables, prior academic performance, and current academic setting in order to bolster the argument that the impact of HwF is not simply a reflection of "better" students completing more activities.

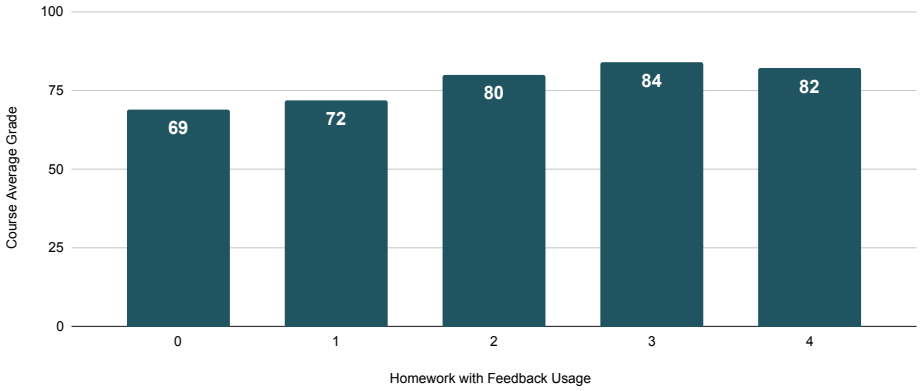
RESULTS

Student-level Analysis

Student-level analyses were focused on student completion patterns - how many HwF activities need to be completed for students to see the benefits? Figure 2 below displays the results of analyzing the impact of HwF usage on final course grade (as a percentage). HwF usage was grouped by students' completion of the median number of HwF activities assigned in their course in order to control for implementation/subject differences. The median was used as an indicator of what was typical or expected within a particular course.

As seen in figure 1, HwF usage category was significantly related to students' final course grade. Students completing a higher percentage of HwF activities also earned higher course grades. Students who completed at the median or greater than the median number of HwF in their course (groups 3 and 4) had an increase of approximately 12-15 grade percentage points compared to students who completed no or less than 50% of the median (groups 0 and 1).

Figure 1. Student Final Grade Performance by HwF Usage Category



Note
The HwF usage categories were defined as follows:

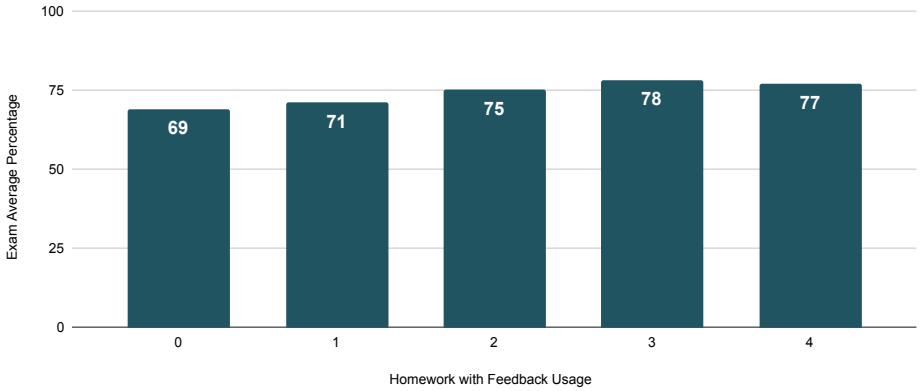
0– No HwF completed
1– Less than 50% of the overall median completed,
2– 50%-99% of the overall median completed,
3– 100% of the overall median completed,
4– Greater than the median.

Many participating study instructors included Achieve activity performance, including HwF completion, into students' final grade calculation. Therefore, it was important to validate the effect of HwF on student learning by examining another outcome. Researchers were given access to students' course exam average (as a percentage) by instructors for consenting students. Since exams were not directly tied to Achieve nor HwF, this would provide evidence that HwF impacted student learning outcomes more generally.

Figure 2 below displays the results of analyzing the impact of HwF usage on exam average (as a percentage). Similar to the course grade analysis, HwF usage was grouped by students' completion of the median number of HwF activities completed in their course.

As seen in the figure, the HwF usage category was significantly related to students' average exam grade. Students completing a higher percentage of HwF activities also earned higher average exam grades. Students who completed at the median or greater than the median number of HwF in their course (groups 3 and 4) had an increase of approximately 7-9 percentage grade points compared to students who completed no or less than 50% of the median (groups 0 and 1).

Figure 2. Student Exam Performance by HwF Usage Category



Note
The HwF usage categories were defined as follows:

0– No HwF completed
1– Less than 50% of the overall median completed,
2– 50%-99% of the overall median completed,
3– 100% of the overall median completed,
4– Greater than the median.

Interaction with First Generation College and Financial Aid Students

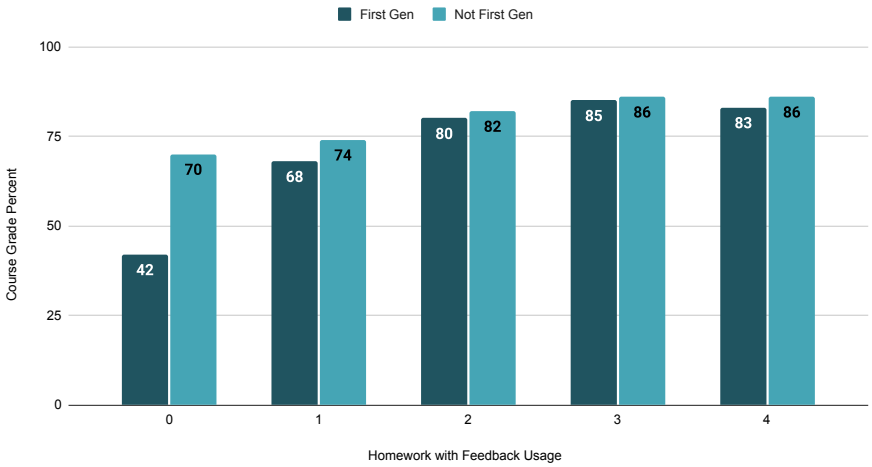
The researchers were also interested in how use of HwF could be particularly beneficial for underrepresented groups and students experiencing barriers to college success. HwF within the Achieve environment may enable students to practice and test their knowledge, increasing their self-efficacy without triggering stereotype threat. To do so, use of HwF was examined as an interaction with gender, race/ethnicity, first generation college student status and eligibility for financial aid to see if HwF had an even greater benefit for certain students. An interaction occurs when the impact of one variable on an outcome depends on another variable.

There was not a significant interaction between HwF usage and gender nor race/ethnicity in predicting student outcomes (course grade), meaning the impact of using HwF was similar across these groups. There was a significant interaction between HwF usage and first generation college student status, meaning the impact of using HwF was different for these students. There was also a significant interaction between HwF usage and eligibility for financial aid, meaning the impact of using HwF was different for these students as well.

Figures 3 and 4 display the results of analyzing the interaction between HwF usage and first generation college student status/financial aid status on final course grade. Increased usage of HwF was particularly beneficial for first generation college students' final grade compared to non-first generation college students, as well as to financial aid recipients compared to non financial aid recipients.

As seen in figure 3, as HwF usage increases, the differences in final grade between first generation students and non-first generation students is diminished. When first generation students are completing no HwF assignments in their course (group 0), their final grade is about 28 percentage grade points lower than their non-first generation peers who similarly completed no assignments. However, when they completed 100% of the median or greater (groups 3 and 4), that difference is no longer statistically significant and reduced to 3 percentage grade points or less.

Figure 3. Course Grade Performance by HwF Usage Category X First Generation Status

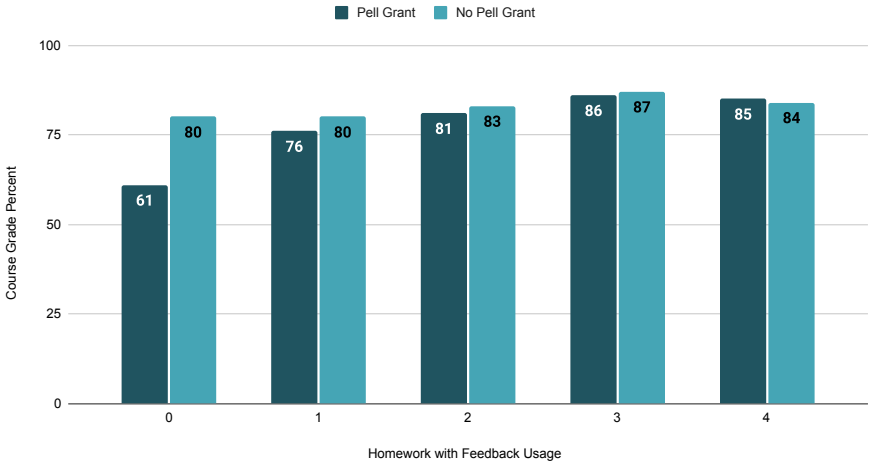


Note
The HwF usage categories were defined as follows:

0– No HwF completed
1– Less than 50% of the overall median completed,
2– 50%-99% of the overall median completed,
3– 100% of the overall median completed,
4– Greater than the median.

This pattern was replicated for students eligible for financial aid. As seen in figure 4, as HwF usage increases, the differences in final grade between financial aid eligible students and non financial aid students is diminished. When financial aid eligible students are completing no HwF assignments in their course (group 0), their final grade is about 19 percentage grade points lower than their non financial aid peers who similarly completed no assignments. However, when they completed 100% of the median or greater (groups 3 and 4), that difference is no longer statistically significant and reduced to 1 percentage grade point.

Figure 4. Course Grade Performance by HwF Usage Category X financial aid Status



Note
The HwF usage categories were defined as follows:

0– No HwF completed
1– Less than 50% of the overall median completed,
2– 50%-99% of the overall median completed,
3– 100% of the overall median completed,
4– Greater than the median.

The reduction in the difference in performance between underrepresented students and their peers is often referred to as bridging the "equity gap", which is critically important so that all learners have similar opportunities to experience success.

IMPLICATIONS FOR INSTRUCTORS

Overall, the research findings suggest a benefit of HwF on student learning and academic outcomes. Students who completed more HwF activities compared to what was typical in their course (median) had higher course grades and exam grades. Students who were the first in their family to go to college, as well as financial aid eligible students, especially benefited from completing more HwF activities.

Instructors have a powerful tool at their disposal in the form of HwF within Achieve, which can be harnessed to facilitate students' mastery of course topics within a supportive learning environment. This approach is particularly valuable because it not only allows students to

practice and refine their understanding but also provides them with constructive and targeted feedback, which is an essential component of the learning process. Additionally, the potential benefits of HwF can be further amplified when it is integrated with other active-learning pedagogy techniques. Using HwF as a part of active learning pedagogy becomes a more dynamic and effective tool for promoting student engagement, comprehension, and growth. It encourages students to take an active role in their learning journey and provides educators with insights into their progress.

Based on the previous results, instructors should assign HwF activities as an opportunity for students to apply what they've learned during instruction, reinforcing their understanding of the subject matter. Students should complete at least what is typical in their course (median) or above to see gains in their course performance. Furthermore, while this was not tested directly as part of the current research study, prior research suggests instructors who use the available information on students' HwF performance in order to identify gaps in students' knowledge, as well as a signal to adapt instruction, provide particular resources, or enact interventions, would likely see an even greater benefit of HwF.

DISCUSSION

The present study investigated the impact of completing HwF in Achieve across a range of institutions, instructors, and students. While making definitive claims about causation requires additional experimental research, the findings suggest a significant association between engaging in HwF activities and students' academic performance. The chance to practice and evaluate their understanding while receiving targeted feedback exhibited a positive relationship on course performance. This connection was particularly notable among students who were the first in their families to attend college and students eligible for financial aid.

Research has shown that self-efficacy plays a significant role in academic performance, surpassing prior knowledge, and it holds particular importance for underrepresented student populations (Ballen et al., 2018; Elias & MacDonald, 2007). The valuable feedback and educational resources provided by HwF could have empowered first-generation and financial aid students to learn from their errors and pinpoint misconceptions, potentially leading to increased self-efficacy in the subjects covered in their courses.

Furthermore, the HwF environment may have been less likely to activate stereotype threat, allowing students to channel their entire cognitive effort toward the task at hand. Ongoing research is needed to investigate the potential mechanisms underpinning this association.

LIMITATIONS AND FUTURE RESEARCH

While the current work represented a large and diverse sample, a convenience sample was used. This was not a true experiment with random assignment. A multitude of variables were used to serve as statistical controls, but the lack of random assignment is a limitation. Individual differences of students not captured by the variables used as controls cannot be ruled out as potential confounding variables.

Instructors' implementation of HwF activities was also not controlled. Instructors were free to use HwF as much or as little as they deemed necessary. Furthermore, instructors' use of low-stakes assessment or active-learning strategies more generally outside of Achieve and HwF was not measured or controlled. Some instructors may incorporate these strategies more or less in their instruction, which may affect the impact of HwF.

Future experimental studies could test the impact of HwF by randomly assigning students within the same course or instructor to either receive HwF activities or not. This design would help strengthen arguments of causality by ruling out both individual differences and instructional differences as potential explanations for group differences. Future research could also incorporate qualitative methods to complement the quantitative analyses. Qualitative methods such as observations of instruction and in-depth interviews with instructors and students could help identify pedagogical patterns including low-stakes assessment/active-learning. A more complete understanding of how instructors are incorporating these strategies into their courses can help clarify how HwF can complement and bolster instruction.

REFERENCES

Ballen, C. J., Wieman, C., Salehi, S., Searle, J. B., & Zamudio, K. R. (2017). Enhancing diversity in undergraduate science: Self-efficacy drives performance gains with active learning. CBE—Life Sciences Education, 16(4), ar56.

Bennett, R. E. (2010). Cognitively based assessment of, for, and as learning (CBAL): A preliminary theory of action for summative and formative assessment. Measurement, 8(2-3), 70-91.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. Assessment in Education: principles, policy & practice, 5(1), 7-74.

Cogliano, M., Bernacki, M. L., & Kardash, C. M. (2021). A metacognitive retrieval practice intervention to improve undergraduates' monitoring and control processes and use of performance feedback for classroom learning. Journal of Educational Psychology, 113(7), 1421.

Cotner, S., & Ballen, C. J. (2017). Can mixed assessment methods make biology classes more equitable?. PLoS One, 12(12), e0189610.

Delaney, P. F., Verkoeijen, P. P., & Spirgel, A. (2010). Spacing and testing effects: A deeply critical, lengthy, and at times discursive review of the literature. Psychology of learning and motivation, 53, 63-147.

Dennehy, T. C., Smith, J. S., Moore, C., & Dasgupta, N. (2018). Stereotype threat and stereotype inoculation for underrepresented students in the first year of college. In R. S. Feldman (Ed.), The first year of college: Research, theory, and practice on improving the student experience and increasing retention (pp. 309–344). Cambridge University Press.

Elias, S. M., & MacDonald, S. (2007). Using past performance, proxy efficacy, and academic self-efficacy to predict college performance. Journal of Applied Social Psychology, 37(11), 2518-2531.

Haak, D. C., HilleRisLambers, J., Pitre, E., & Freeman, S. (2011). Increased structure and active learning reduce the achievement gap in introductory biology. Science, 332(6034), 1213-1216.

Han, J. H., & Finkelstein, A. (2013). Understanding the effects of professors' pedagogical development with Clicker Assessment and Feedback technologies and the impact on students' engagement and learning in higher education. Computers & Education, 65, 64-76.

John-Henderson, N. A., Rheinschmidt, M. L., Mendoza-Denton, R., & Francis, D. D. (2014). Performance and inflammation outcomes are predicted by different facets of SES under stereotype threat. Social Psychological and Personality Science, 5(3), 301-309.

Malespina, A., & Singh, C. (2022). Gender differences in test anxiety and self-efficacy: why instructors should emphasize low-stakes formative assessments in physics courses. European Journal of Physics, 43(3), 035701.

Roediger III, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. Psychological science, 17(3), 249-255.

Schmader, T. (2010). Stereotype threat deconstructed. Current Directions in Psychological Science, 19(1), 14-18.

Shute, V. J. (2008). Focus on formative feedback. Review of educational research, 78(1), 153-189.

Thomas, A., & Cross, P. "Minute paper." Classroom assessment techniques: A handbook for college teachers, 2, 148-153.