# Identifying Students Struggling in Courses by Analyzing Exam Grades, Self-reported Measures and Study Activities

Christensen, Bianca Clavio; Bemman, Brian; Knoche, Hendrik; Gade, Rikke

*Published in:*
The Interplay of Data, Technology, Place and People for Smart Learning

*DOI (link to publication from Publisher):*
10.1007/978-3-319-92022-1_15

*Publication date:*
2018

*Document Version*
Accepted author manuscript, peer reviewed version

# Identifying Students Struggling in Courses by Analyzing Exam Grades, Self-reported Measures and Study Activities

Bianca Clavio Christensen, Brian Bemman, Hendrik Knoche, and Rikke Gade

Aalborg University, Rendsburggade 14, 9000 Aalborg, Denmark,
{bcch|bb|hk|rg}@create.aau.dk

**Abstract.** Technical educations often experience poor student performance and consequently high rates of attrition. Providing students with early feedback on their learning progress can assist students in self-study activities or in their decision-making process regarding a change in educational direction. In this paper, we present a set of instruments designed to identify at-risk undergraduate students in a Problem-based Learning (PBL) university, using an introductory programming course between two campus locations as a case study. Collectively, these instruments form the basis of a proposed learning ecosystem designed to identify struggling students by predicting their final exam grades in this course. We implemented this ecosystem at one of the two campus locations and analyzed how well the obtained data predicted the final exam grades compared to the other campus, where midterm exam grades alone were used in the prediction model. Results of a multiple linear regression model found several significant assessment predictors related to how often students attempted self-guided course assignments and their self-reported programming experience, among others.

**Keywords:** Academic performance, Student retention, Learning Management System, Learning Tools Interoperability, Problem-Based Learning, Flipped learning

## 1 Introduction

Students enrolled in educations with technical content often struggle with passing technical courses and frequently drop out as a result [9, 4]. Much of the research on student dropouts or retention has focused on the personality traits of a student, typically without also considering their actual learning progress. Both struggling students and course instructors, however, can benefit from an understanding of how the learning process of students is progressing. Such an understanding, for example, might encourage students to engage more deeply with the learning material or allow instructors to better direct resources to those in need. With many openly accessible learning resources, such as *Massive Open Online Courses (MOOCs)* now available, teachers are able to construct diverse learning ecosystems for their students which extend far beyond institutionally

managed, digital *Learning Management Systems (LMSs)*. The research question addressed here is how information gathered from the diverse interactions students have with these learning resources can be used to identify struggling individuals.

In this paper, we present a set of instruments designed to identify struggling first-semester undergraduate students enrolled in an introductory programming course at a *Problem-based Learning (PBL)* university in Denmark, named Aalborg University (AAU). These instruments consist of both student self-reported personal attributes and self-assessed measures of the learning progress. We use these instruments in the construction of a multiple linear regression model for predicting student final exam scores. Our proposed model consists of significant predictors from this set of instruments that suggest a possible relationship between the academic success of a student and select personal attributes and learning progress measures. Importantly, these predictors could provide the university with the means to more effectively identify struggling students who may be at risk of leaving the education, allowing them to offer guidance to these individuals as early in their education as possible.

## 2 Background

Previous research on student retention has identified a number of factors for decreasing the risk of students leaving educational programs: *growth mindset* [11], *grit* (i.e., perseverance when faced with challenges) [10], *study habits* [27], *high school habits* [15, 19], and *social support for studying* [5]. Although this research has documented a wide range of potential predictors of student retention, agreement between studies is low [18, 20, 26]. For this reason, continued research would be better served by considering case studies [9]. This could be done, for example, by detecting students at risk of leaving the education and then directing adequate resources to those individuals based on relevant features of the study program from which these students left. One previous study on student dropouts [4], looked at first semester students in an undergraduate Media Technology (MEA) program at AAU. These findings provided some evidence, in the form of interviews and study diary logs, that suggested that the skills in mathematics and programming required of the program were higher than students initially expected, resulting in a high number of dropouts. Natural science courses, for example, are notorious for low retention and first-time success rates, particularly in the first year of study [25]. It is essential then to investigate interdisciplinary educations, such as MEA, that combine technical, scientific, and design skills.

Engaging students in the learning process and making them responsible for their own progress is one of the primary goals of university education and PBL, in particular. One important reason for this is the comparatively less interaction and feedback students receive at a university than they are used to in high school. The principle of *pre-training* [23] suggests that cognitive overload can be reduced by providing students with basic information ahead of their actual lectures. This principle is often implemented online as self-study activities with *flipped learn-*

*ing*. Such an approach leaves more time for the instructor to facilitate classroom activities that are essential to a PBL framework, where students are required to analyze, evaluate, and create content in a hands-on fashion [17]. These activities in PBL can include scaffolding more complex concepts and skills through interaction, group work, peer feedback, and immediate teacher support [16]. This stands in contrast to, for example, the earlier stages in Bloom's taxonomy of learning (remember, understand, and apply) [3].

Breaking material down into smaller parts is one way to reduce the cognitive load of students [1] and doing so can make that content more accessible, focused, and easier to digest. Self-assessment questions are one way to increase this sort of in-depth learning in students [12], e.g., when a student is trying to understand where they went wrong on a quiz. Through self-assessment questions, instructors can efficiently assess and manage student learning by creating, for example, online assignments with automated grading and feedback. This is a fundamental approach advocated in *Learning Tools Interoperability (LTI)* [21]. Moreover, self-assessment quizzes of this type can be designed to adapt and grow in response to student performance based on, for example, the previous answers they provide to the system. However, creating such content is time-consuming and provides little personal control for the teacher when implemented in an LMS, such as Moodle [21, 24]. This ability to adjust feedback to a student's zone of *proximal development* [28] is often considered the gold standard of education that current digital tools and systems, unfortunately, do not meet. Even so, more and more teachers are relying on online activities for instruction. An important feature of this approach is the teacher's ability to not only monitor student progress but also target struggling students for early intervention. Doing this online also allows immediate response and communication between teacher and students when adjusting instructions, moderating difficult learning content, and addressing student misunderstandings [2]. Additionally, an instructor can get an idea of a student's level of engagement in a course by observing the relationship between that student's performance and their use of Moodle [6]. However, the relationship between grades and behavioral data such as Moodle activity logs is complex and influenced by additional factors [6].

## 3 Case study context and the learning ecosystem

AAU operates according to a PBL model which assumes that students learn best when applying theory and research-based knowledge to collaborative working strategies aimed at real-world problems. In any one educational program at AAU, each student must enroll in semester study activities corresponding to a total workload of 30 ECTS credits, where a single ECTS is anywhere between 25 to 30 work hours. These 30 ECTS credits typically include a semester project worth 15 ECTS and three courses worth 5 ECTS each. The mandatory study activities at AAU (i.e., semester projects and courses) require students to make connections between them that span from course-to-course in a single semester as well as across semesters.

In the MEA program, the introductory programming course required in the first semester constitutes an important building block for a student's academic success in further semesters. The programming course is offered at two campus locations, one in Aalborg (AAL) and one in Copenhagen (CPH). While the course is primarily held for MEA students, in AAL a group of non-technical Product and Design Psychology students ($AAL_P$) must also enroll as part of their degree program. At both locations, instructors make an effort to "harmonize" their shared syllabus and content, however, their teaching methods may differ.

At AAL, the instructor integrated several methods of flipped instruction [16] in the Fall semester 2017 that were not used in CPH. These methods include online self-study activities consisting of self-assessment quizzes (SA), exercises on khanacademy.org (KA), and mandatory hand-in assignments on peergrade.io (PG). A Moodle course page served as the LMS for providing access to these self-study activities in addition to a collection of other learning resources, such as supplementary video content. The course utilized a combination of online instructions and face-to-face lessons, and when combined with these self-study activities, formed a learning ecosystem which encourages student learning beyond the boundaries of the classroom [7, 13, 14]. Fig. 1 shows an overview of this learning ecosystem used by AAL in its introductory programming course.

This learning ecosystem consisted of several online self-study activities, including course readings, videos, SA and KA, which encourage student learning prior to class. During lectures, the teacher presented the topic in a shortened format, followed by hands-on programming exercises in which the students may either work alone or in groups with the help of the teacher, teaching assistants, or their peers as part of study cafes. These in-class opportunities were designed to reinforce the concepts learned in the self-study activities through practical

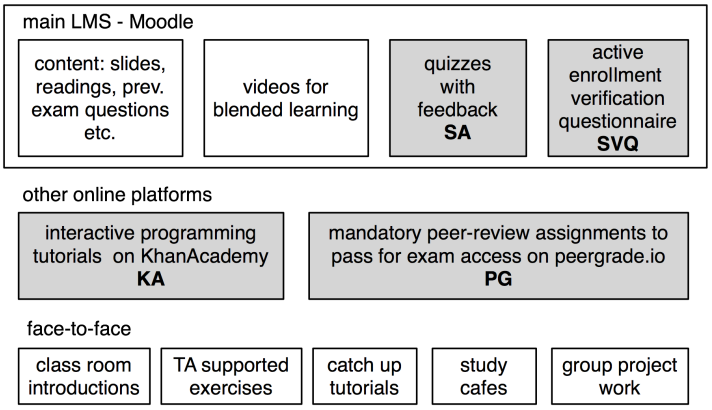



Fig. 1: Overview of the learning ecosystem used by AAL campus location in its introductory programming course in the Fall semester 2017.

experience. At various points in the semester, students were asked to complete programming assignments and evaluate those created by their peers in PG. Such peer learning provided students with the opportunity to critically apply their knowledge.

MEA students at AAU have diverse backgrounds (e.g., in nationality, high school specialization, and proficiency in math) and study interests (e.g., in design or programming) [4]. In order to gather information about this diversity in first-year AAL students useful as control variables, we designed a survey called the *Study Verification Questionnaire (SVQ)*. It consisted of a set of 111 self-reported questions based on established factors for student retention discussed in Section 2, such as grit and study habits, among others.

## 4 Data collection and method

In order to discover possible assessment predictors for identifying struggling students, we gathered data from all AAL students for each of the self-assessment activities shown in grey in Fig. 1. Additionally, we collected the students' scores from MT and the final exam from both AAL and CPH. In Section 5, we explore the use of these assessments, SVQ, SA, KA, PG, and MT, in the construction of a multiple linear regression model for predicting final exam scores.

Students from $\text{AAL}_M$, $\text{AAL}_P$, and $\text{CPH}_M$ all took the same MT and the same final exam. However, not all students who completed MT took the final exam. Some students in $\text{AAL}_M$, for example, were ineligible to attend the final exam due to unfulfilled course requirements, e.g., not completing PG hand-in assignments. Following MT, we invited AAL students with the lowest MT scores to attend tutoring sessions as a first step towards providing early and targeted academic intervention. Unfortunately, only a handful of students attended one or more of these sessions. Table 1 shows the number of students in AAL (both educations) and CPH who completed the MT and attended tutoring, as well as the average final exam scores for these students.

Table 1: Number of students in the AAL and CPH campus locations and their average scores on the midterm (MT) and final exams.

| | | $\text{AAL}_M$ | $\text{AAL}_P$ | $\text{CPH}_M$ | Total |
|---|---|---|---|---|---|
| MT | Avg. score | 66.67 | 73.72 | 63.65 | 65.78 |
| | Failed students | 40 | 4 | 36 | 80 |
| | Total students | 84 | 21 | 82 | 187 |
| Tutoring | Not invited | 36 | 17 | - | 53 |
| | Invited | 48 | 4 | - | 52 |
| | Attended | 7 | 2 | - | 9 |
| Exam | Avg. score | 44.00 | 47.86 | 43.96 | 44.66 |
| | Failed students | 31 | 5 | 29 | 65 |
| | Total students | 72 | 22 | 83 | 177 |

We began the analysis of our assessment data by investigating how well MT alone could predict the final exam scores of students at both AAL and CPH. From here, we explored additional predictors related to student engagement and learning progress (i.e., SA, KA, and PG) found in the learning ecosystem implemented at AAL. We concluded our analysis by seeing how well self-reported measures in SVQ could improve the prediction of final exam scores at AAL.

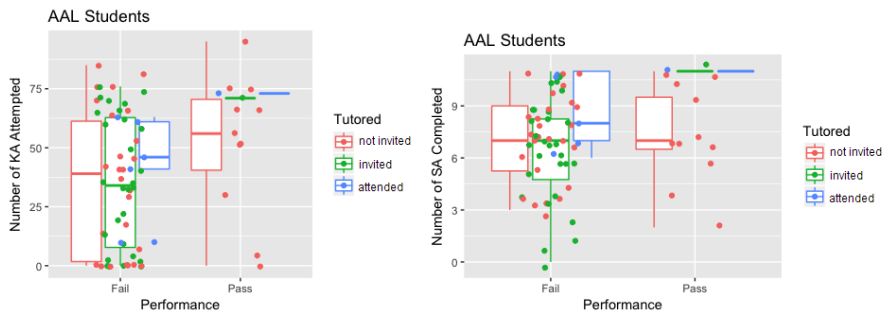
## 5 Results

In order to predict the final exam scores, we began by using the students' scores on MT from both AAL and CPH to construct two base linear regression models. The results from these initial models were rather poor when using adjusted r-squared as the measure of performance (AAL $r^2 = 0.50, p < 0.001$, CPH $r^2 = 0.46, p < 0.001$). However, due to the learning ecosystem implemented by the AAL campus location, we were able to test the significance of including the additional assessments used by AAL and compare their performance to the base models for both AAL and CPH. From the AAL base model, we constructed a multiple linear regression model by selecting from our additional assessments, SA, KA, and PG. The selection of terms in this model was done using a bi-directional step-wise method using AIC as the selection discriminator. During this process, we used the number of SA assignments completed, the number of KA attempts, and the PG score. Although both SA scores (and not number of attempts or completed assignments) and PG scores proved to be significant alone, they did not improve the overall model using our chosen method for model selection. The best model from this selection method was MT + KA ($r^2 = 0.53, p < 0.001$). Fig. 2 shows a comparison of the significant predictor KA and the insignificant predictor SA used in our base model for AAL students who did and did not pass the final exam. Fig. 3 (a) shows the insignificant predictor PG in our base model for AAL students who did and did not pass the final exam.

Before exploring the role of all SVQ questions in our model, we considered that because the focus of the course was programming, it would be worth seeing the effect that self-reported programming experience ($\text{SVQ}_{pe}$) had on the improved model of MT + KA. We began by manually selecting from SVQ the two questions related to student self-reported programming experience and averaged their values. The addition of this predictor to MT + KA resulted in a marginal improvement ($r^2 = 0.55, p < 0.001$). Fig. 3 (b) shows the significant predictor $\text{SVQ}_{pe}$ used in our improved model for AAL students who did and did not pass the final exam.

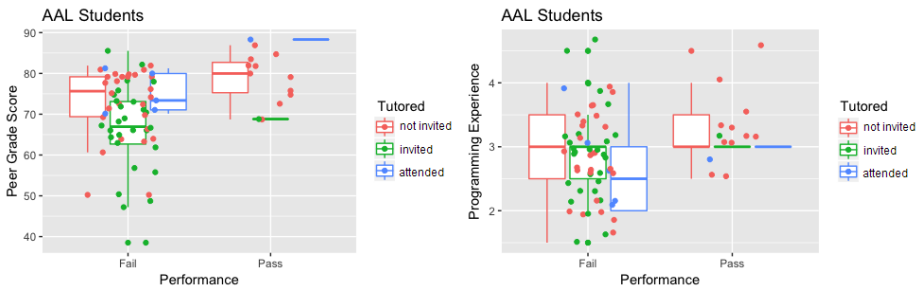
In our approach, we have elected to manually select individual predictors from SVQ that improve our model of MT + KA + $\text{SVQ}_{pe}$ using adjusted r-squared as a selection criteria. We subsequently dropped each predictor after testing so that at any one time, only a single predictor is used. Following this procedure, we arrived at five SVQ predictors that together significantly improved our model ($\text{SVQ}_{subset}$). These predictors each belong to one of five

different categories related to a student's behavior and psychology. These categories were *personal trait comparison* [19, 15], *growth mindset* [11], *high school trust* [29, 8], *attitude towards education* [22], and *understanding of the Media Technology education* [4]. Our final model and best fitting linear regression, $\text{AAL}_{final}$, consists of the following assessment predictors: MT + KA + $\text{SVQ}_{pe}$ + $\text{SVQ}_{subset}$ ($r^2 = 0.72, p < 0.001$). For comparison, an ANOVA test revealed a significant difference from our AAL base model shown above and $\text{AAL}_{final}$ ($F = 9.32, p < 0.001$).




(a) (b)

Fig. 2: Box plots grouped according to AAL student performance (pass or fail) on the final exam and tutoring status. Plots show the spread of number of attempted KA assignments in (a) and the number of completed SA assignments in (b).




(a) (b)

Fig. 3: Box plots grouped according to AAL student performance (pass or fail) on the final exam and tutoring status. Plots show the spread of PG scores in (a) and $\text{SVQ}_{pe}$ in (b).

## 6 Discussion

While the performances of the two base linear models for AAL and CPH were roughly the same, our final AAL model, containing the assessment predictors MT + KA + $\text{SVQ}_{pe}$ + $\text{SVQ}_{subset}$, proved significantly better than the base CPH model (AAL $r^2 = 0.72, p < 0.001$, CPH $r^2 = 0.46, p < 0.001$). When building our final model, the improvement to the AAL base model by KA and not, for example, by SA, as shown in Fig. 2 could be due to the greater number of possible assignments in the former (103 KA assignments and 42 SA assignments). Even considering their difference in number, students attempted a proportionally greater number of KA assignments than SA.

There are likely a number of reasons for why SA and PG were not selected. With SA, students are allowed to re-take each exercise as many times as they would like. This might encourage "high-score seeking" behavior over actual knowledge retention. There was also a significant drop off in the number of completed SA assignments following the midterm. While PG marginally improved the model, it was not significant given our chosen method for model selection. This could be due to peer assessments being incomparable with one another, whereas having only one reviewer (e.g., a teacher) might ensure a more similar assessment across assignments.

It should be noted that in initial tests using all of SVQ, 40 of the total 111 questions were found to be linearly dependent. This finding of multicollinearity suggests a high degree of redundancy in the questions that could be improved in future uses of SVQ. Additionally, our number of predictor variables in SVQ far exceeds the number of observations. For these reasons, a step-wise selection model may not be wholly appropriate. An alternative method of model selection such as Lasso or Principal Component Analysis (PCA) would likely be more robust to such situations.

Overall, our results indicate that while student performance on MT was positively correlated with performance on the final exam in the AAL and CPH base models, having a learning ecosystem which consists of several appropriate and diverse assessments, as demonstrated by the final AAL model, significantly improved the prediction of final exam scores.

## 7 Conclusion

The instruments of the learning ecosystem presented in this paper provide initial findings in support of additional strategies for targeting struggling students in a PBL environment. While the results leave much room for improvement, they nonetheless demonstrate that regular student feedback through self-regulated knowledge assessments, targeted tutoring, and proper evaluations of student behavior and psychology may be essential factors in reducing rates of PBL student failure. Moreover, technological learning tools through, e.g., Moodle, Peergrade or Khan Academy, might serve as useful tools for ensuring the academic success of students. Such benefits are particularly needed at universities where more

and more degree programs are becoming interdisciplinary and courses are being taught by different instructors at separate campus locations. Guaranteeing the quality of education in these situations is essential.

## 8 Future work

In the future, we hope to implement the significant assessments of our learning ecosystem into a system for identifying struggling students prior to the midterm of a given course and incorporate additional sources of relevant information such as Moodle course activity. The variety of significant assessment predictors in our final model emphasizes the need for a learning ecosystem that is both targeted and wide-ranging, as shown, for example, in Fig. 1. With this model, it might be possible in future semesters to target struggling students even before the start of a course by identifying those who reported a low $\mathrm{SVQ}_{pe}$ score. In compliance with LTI, an automated weekly analysis could invite students in need to group tutoring sessions based on how often students either attempted KA or completed SA assignments.

## References


1. Banas, J. R., Velez-Solic, A.: Designing Effective Online Instructor Training and Professional Development. Virtual Mentoring for Teachers: Online Professional Development Practices: Online Professional Development Practices, pp. 1 (2012).
2. Bergmann, J., Sams, A.: Flipping for Mastery. Educational Leadership **71**, pp 24–29 (2014).
3. Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., Krathwohl, D. R. (1956): Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain. New York: David McKay Company.
4. Bøgelund, P., Justesen, K., Kolmos, A., Bylov, S. M.: Undersøgelse af frafald og fastholdelse ved medialogi og andre uddannelser ved det Teknisk-Naturvidenskabelige Fakultet 2015-2016: Arbejdsrapport Nr. 1. Aalborg Universitet (2016).
5. Cabrera, A. F., Nora, A., Castaneda, M. B.: College Persistence: Structural Equations Modeling Test of an Integrated Model of Student Retention. The Journal of Higher Education **64**(2), pp. 123–139 (1993).
6. Casey, K., Gibson, P., Paris, I.: Mining moodle to understand student behaviour. In International Conference on Engaging Pedagogy (2010).
7. Chang, M., Li, Y.: Smart Learning Environments. Lecture Notes in Educational Technology. Springer-Verlag (2015)
8. Cohen, G. L., Garcia, J., Purdie-Vaughns, V., Apfel, N., Brzustoski, P.: Recursive Processes in Self-Affirmation: Intervening to Close the Minority Achievement Gap. Science **324**(5925), pp. 400–403 (2009).
9. Dekker, G., Pechenizkiy, M., Vleeshouwers, J.: Predicting Students Drop out: A Case Study. Proc. of Educational Data Mining (2009).
10. Duckworth, A. L., Quinn, P. D.: Development and Validation of the Short Grit Scale (Grit–S). Journal of Personality Assessment **91**(2), pp. 166–174 (2009).

11. Dweck, C. S.: Self-theories: Their Role in Motivation, Personality, and Development. Psychology Press (2000).
12. Evans, C., Palacios, L.: Interactive self assessment Questions within a Virtual Environment. International Journal of E-Adoption (IJEA) **3**, pp. 1–10 (2011).
13. Giovannella, C.: Smart Learning Eco-Systems: "fashion" or "beef"?.Journal of e-Learning and Knowledge Society, 10(3) (2014)
14. Giovannella, C., Rehm, M.: A critical approach to ICT to support participatory development of people centered smart learning ecosystems and territories. Aarhus Series on Human Centered Computing, 1(1), 2 (2015)
15. Glynn, J. G., Sauer, P. L., Miller, T. E.: A Logistic Regression Model For The Enhancement Of Student Retention: The Identification Of At-Risk Freshmen. International Business & Economics Research Journal **1**(8) (2011).
16. Green, L. S., Banas, J., Perkins, R.: The Flipped College Classroom: Conceptualized and Re-Conceptualized. Springer (2016).
17. Guerrero, W.: Flipped Classroom and Problem-Based Learning in Higher Education. Latin-American Context. Conference Proceedings. The Future of Education, pp. 118 (2017).
18. Herzog, S.: Measuring determinants of student return vs. dropout/stopout vs. transfer: A first-to-second year analysis of new freshmen. In Proc. of 44th Annual Forum of the Association for Institutional Research (AIR) (2004).
19. Higher Education Research Institut: CIRP Freshman Survey – HERI, Retrieved from: https://heri.ucla.edu/cirp-freshman-survey/ (2017).
20. Lassibille, G., Gomez, L. N.: Why do higher education students drop out? Evidence from Spain, Education Economics **16**(1), pp. 89-105 (2007).
21. Leal, J. P., Queirós, R.: Using the learning tools interoperability framework for LMS integration in service oriented architectures. Technology Enhanced Learning TECH-EDUCATION'11 (2011).
22. Levitz, R. N.: Retention Management System Plus Samples Survey and Report Samples, https://www.ruffalonl.com/complete-enrollment-management/student-success/rnl-retention-management-system-plus/samples (2012).
23. Mayer, R. E.: Principles for managing essential processing in multimedia learning: Segmenting, pretraining, and modality principles. The Cambridge handbook of multimedia learning, pp. 169–182 (2005).
24. Severance, C., Hardin, J., Whyte, A.: The coming functionality mash-up in Personal Learning Environments. Interactive Learning Environments **16**(1), pp. 47-62 (2008).
25. Seymour, E., Hewitt N.: Talking about leaving. Boulder: Westview Press (1997).
26. Touron, J.: The determination of factors related to academic achievement in the university: implications for the selection and counseling of students, Higher Education 12, pp. 399-410 (1983).
27. Tsukayama, E., Duckworth, A. L., Kim, B.: Domain-specific impulsivity in school-age children. Developmental Science **16**(6), pp. 879–893 (2013).
28. Vygotsky, L.: Zone of proximal development. Mind in society: The development of higher psychological processes, pp. 157 (1987).
29. Yeager, D. S, Purdie-Vaughns, V., Garcia, J., Apfel, N., Brzustoski, P., Master, A., Hessert, W. T., Williams, M. E., Cohen, G. L.: Breaking the cycle of mistrust: Wise interventions to provide critical feedback across the racial divide. Journal of Experimental Psychology: General **143**(2), pp. 804–824 (2014).