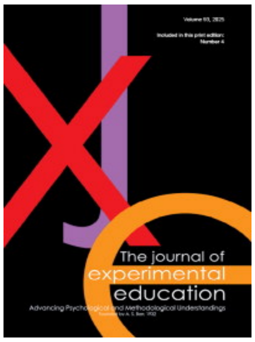

The Journal of Experimental Education

ISSN: 0022-0973 (Print) 1940-0683 (Online) Journal homepage: www.tandfonline.com/journals/vjxe20

Supporting Identity-Based Motivation: Next-Year Continuation Effects of High-Fidelity *Pathways-to-Success*

Daphna Oyserman, Mark Lai, Alysia Burbidge & Nicholas Sorensen

To cite this article: Daphna Oyserman, Mark Lai, Alysia Burbidge & Nicholas Sorensen (08 Oct 2025): Supporting Identity-Based Motivation: Next-Year Continuation Effects of High-Fidelity *Pathways-to-Success*, The Journal of Experimental Education, DOI: 10.1080/00220973.2025.2565170

To link to this article: https://doi.org/10.1080/00220973.2025.2565170

THE JOURNAL OF EXPERIMENTAL EDUCATION
https://doi.org/10.1080/00220973.2025.2565170

MOTIVATION AND SOCIAL PROCESSES
OPEN ACCESS
Check for updates

Supporting Identity-Based Motivation: Next-Year Continuation Effects of High-Fidelity *Pathways-to-Success*

Daphna Oyserman[a], Mark Lai[a], Alysia Burbidge[a], and Nicholas Sorensen[b]

[a]Psychology, University of Southern California, Los Angeles, CA, USA; [b]American Institutes for Research, Chicago, IL, USA

ABSTRACT

We document the next-year carryover effects on grades and risk of course failure ($N = 875$) of *Pathways-to-Success*, a whole-class, teacher-trained, and teacher-led universal intervention. Structural equation modeling reveals that higher fidelity is associated with higher identity-based motivation. Through identity-based motivation, higher fidelity is associated with better 8th-grade academic outcomes, and through 8th-grade outcomes, 9th-grade outcomes. Providing support for the theorized process model, effects are robust across tools for coding and modeling the identity-based motivation construct. *Pathways-to-Success* supports school-focused identity-based motivation and hence, academic attainment, *via* activities that bolster students' existing identity-based motivation in three ways. Activities help students see that they and their peers all (1) have skills and abilities to succeed in school this year and into the future, (2) imagine some future as an adult with school as the path to get there, and (3) interpret difficulties and setbacks as indicators of importance, not only of low odds or impossibility. Our results contribute to studies documenting the effects of *Pathways-to-Success* on 8th-grade outcomes by extending effects to 9th-grade and suggesting the robustness of the key mediating construct, increasing confidence in the effects and process model.

KEYWORDS
Identity-based motivation; social and behavioral intervention; sustainability; teacher-led; transition to high school

Some but not all promising social and behavioral interventions yield discernible effects on student academic outcomes (e.g., Bradshaw et al., 2009; Cook et al., 2014; Gresham, 2004). The successful ones are associated with small but statistically significant improvements in academic outcomes, typically measured at the end of the intervention (e.g., Durlak et al., 2022; Gage et al., 2017; Kase et al., 2017). Successful interventions are more likely to be replicable and hence more worth testing for their staying power and robustness if they are manualized, provide training materials, document implementation fidelity, and test process models (e.g., Dekker & Meeter, 2022; Giner-Sorolla et al., 2024; Zhang et al., 2023). Documenting staying power matters since staying power implies that investing in the intervention will yield longer-term effects. The current paper focuses on a manualized successful intervention, *Pathways-to-Success* (Oyserman, 2015), addressing two related questions: The staying power of impact on academic outcomes from the end of middle school to the first year of high school, and the robustness of the theoretical process model underlying the intervention to two alternate coding

CONTACT Daphna Oyserman oyserman@usc.edu Mind and Society Center, University of Southern California, 635 Downey Way, Verna & Peter Dauterive Hall, Suite 205, Los Angeles, CA 90089-3333 USA.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/00220973.2025.2565170.

and analytic strategies. In the case of the *Pathways-to-Success* intervention, staying power entails effects lasting from middle school through the first year of high school, which is a critical window for preventing student dropout (e.g., Allensworth & Easton, 2007). As detailed next, coding and analytic strategy robustness are important for two reasons. First, they are components of intervention durability. Second, they advance understanding of the posited active ingredient, identity-based motivation (IBM), and hence, research on the future self.

Why consider pathways-to-success

IBM: the theoretical underpinning of pathways-to-success

IBM theory is a social cognitive theory of self, self-regulation, and goal pursuit (Oyserman, 2007, 2009, 2024). It provides a novel perspective on a familiar paradox: people can imagine their future selves without taking or sustaining future-focused action. IBM theory explains how this can be true by focusing on the inferences people draw from difficulty as an underappreciated aspect of how and when identities motivate current action (Elmore & Oyserman, 2012; Oyserman, 2007, 2024; Oyserman et al., 2012). This contrasts with other theories about the future self, which do not consider the inferences people draw from their metacognitive experience of difficulty while thinking about tasks and goals. Instead, these alternative theories focus on the accessibility, valence, or content of the future or possible self, the experienced continuity between the current and future self, or the experienced gap between the current and future self (for a full review, Oyserman & Horowitz, 2023). As reviewed fully by Oyserman and Horowitz (2023), the literature is siloed. Some studies test whether the content, valence, or accessibility of a possible or future self is associated with behavior. Another set of studies tests whether experiencing continuity between the current and future self is associated with behavior. A third set of studies tests whether experiencing a gap between the current and future self is associated with behavior. Each set of studies yields supporting evidence, yet if what matters for behavior is the content of a future imagined identity, why would having experienced continuity between current and future "me" without a focus on identity content matter? And the reverse, if what matters is a general sense of continuity between current and future "me", why would the content or valence of a particular possible identity matter? Parallelly, if what matters for behavior is experienced continuity between current and future "me", why would experiencing a gap matter? Each of these alternatives is supported by some evidence, but none of these alternative models can explain the existence of evidence supporting the alternative models. In contrast, the IBM theory can explain all of it. It does so by focusing on three aspects of motivation and their recursive connections. These three aspects are the inferences people draw about who they are and can become when thinking about tasks and goals feels hard, the dynamic construction of identity content and certainty, and the extent to which actions to attain goals feel like 'me' things to do (for a review, Oyserman, 2007, 2024).

IBM theory predicts that though people prefer to act in ways that fit their future self, the on-the-mind aspect of their future self, what they infer when thinking about that future self, or self-related tasks, feels hard to do, and their in-the-moment actions are recursively related. Specifically, IBM theory posits that when thinking about a task or goal feels hard to think about, people can draw two inferences from that metacognitive experience of difficulty. They can infer it is not for them (difficulty-as-impossibility) and that it is important to them (difficulty-as-importance). If the former is accessible, people tend to feel unsure that they can attain whatever identity is on their minds and withdraw effort from related tasks[1]. Parallelly, feeling less certain that a current or possible identity is a good fit for oneself increases the likelihood that one will make a difficulty-as-impossibility inference when making sense of metacognitive difficulty. In contrast, if a difficulty-as-importance inference is accessible, people tend to feel surer about attaining the on-the-mind identity and continue engaging.

Evidence for these posited processes comes from measurement and experimental data (for a summary, Oyserman, 2024). Although mostly tested in the U.S., the IBM theoretical model is also used to make sense of educational progress outside the U.S. (e.g., in Australia, see Bennett et al., 2016; in Greece, see Efthymiadou et al., 2025). For example, in Greece, how much students endorse difficulty-as-importance and difficulty-as-impossibility, their strategy plausibility score, and having school-focused possible identities have separate associations with grade point average (Efthymiadou et al. 2025). Lee (2024) also documents associations over time with separate effects of difficulty-as-importance and difficulty-as-impossibility scores on probationary college students' positive self-regard and self-regulation.

Turning to tests of the three aspects of IBM, the first aspect entails identity content. As highlighted by the above-noted review by Oyserman and Horowitz (2023), having a possible future identity is necessary but insufficient to motivate action consistently. The second aspect entails the inferences people draw from their metacognitive experiences of difficulty while considering their possible identity. Validated and reliable brief scales for measuring individual differences in endorsing difficulty-as-importance and difficulty-as-impossibility exist (Fisher & Oyserman, 2017; Oyserman, 2024). Discriminant validity studies show that the correlation between how people score on difficulty-as-importance and difficulty-as-impossibility scales is, on average, small and negative, and the confidence interval around the correlation typically includes 0 (the scale scores are not correlated, e.g., Fisher & Oyserman, 2017; Yan et al., 2024). Scale scores are also distinct from related scales including optimism ("Things will be OK"), conscientiousness ("I follow through with my plans"), efficacy ("If I try, I can succeed"), growth mindset ("If I try, I can change"), grit ("I am someone who keeps trying"), locus of control ("My success is in my control"), promotion focus ("I am focusing on attaining successes.") and prevention focus ("I am focusing on avoiding failures"). Rather than answering the question "Will effort matter for me?" or the related question "What is the way for me to strive?", difficulty-as-importance answers the question "Why should I try?" by affirming "Because this matters for me." Difficulty-as-impossibility answers a different question: "Is this where I should put my time and effort?" Making progress on goals requires flexible access to both difficulty-as-importance and difficulty-as-impossibility inferences to focus engagement and reduce perseverance (O'Donnell & Oyserman, 2023; Oyserman, 2024). Intervention studies measuring both inferences among children typically use the six-item, 5-point Likert response (1 = *strongly disagree* to 5 = *strongly agree*, from Elmore et al., 2016).

The third aspect of IBM entails a readiness to act in ways that make sense given accessible identities and inferences from metacognitive difficulty. Beyond general readiness, IBM research documents that goal progress requires an accessible array of strategies. Without strategies linked to accessible identities, aspirations may fail to trigger action over time (for comparison by socioeconomic status, Oyserman et al., 2011). Empirically, sustained action is more likely when strategies are concrete (depicting what the action specifically entails) and account for the social context in which goal striving occurs (anticipating strategies needed given social context, for a review, Oyserman et al., 2012). This combination of having possible identities linked to concrete strategies that connect to the social context has been termed having a "plausible roadmap" or simply "plausibility" (Oyserman et al., 2004). Oyserman and colleagues (2004) had skilled coders code idiographic responses to open-ended prompts regarding next-year expected and to-be-avoided possible identities and actions to work on them to obtain a single plausibility score. The plausibility score predicted academic outcomes better than the content of possible identities alone or other ways of coding possible identities and counting strategies (Oyserman et al., 2004).

However, skilled coding is impossible to sustain at scale. Researchers have explored alternative approaches using large language models (LLMs) to obtain word embeddings that are then transformed into continuous values representing the functional relationship between the idiographic content of possible identities and strategies and GPA using Support Vector Regression (SVR; O'Donnell & Oyserman, 2023). Using such LLM methods rather than a prompt-based LLM like ChatGPT avoids the limitations of content or "natural language" generation (e.g., hallucinations, Ji et al., 2023; Rawte et al., 2023). But, compared with human coding, it does not specifically classify the components of plausibility using previously human-coded data as a benchmark. In the current paper, we address this limitation of prior LLM approaches.

Pathways-to-Success is feasible to implement and yields replicable results

Pathways-to-Success is a social and behavioral intervention program rooted in IBM theory (Oyserman, 2015). It is a whole-class, universal program that translates IBM theory into brief, classroom-period-length activities teachers implement in the first six weeks of the school year (Oyserman, 2015). When implemented as intended, *Pathways-to-Success* supports the three aspects of student IBM: their possible identities, their strategies to work on their possible identities, and the meaning they make of difficulties along the way. Teachers can train other teachers to deliver the *Pathways-to-Success* intervention with fidelity (Horowitz et al., 2018). Successful delivery has proved transferable to different contexts (as noted in Dekker & Meeter, 2022). Classrooms within the U.S. (Oyserman et al., 2002) and outside (e.g., Florea & Negru-Subtirica, 2025) report successful implementation. Fidelity assessment is manualized (Oyserman, 2015), and fidelity matters for student academic outcomes (Oyserman et al., 2021).

The *Pathways-to-Success* implementation manual and fidelity coding measures are published (Oyserman, 2015). The published manual has been commended for reporting the full information needed for replication (e.g., Lewis & Wai, 2021; Premachandra & Lewis, 2022). Google Scholar lists over 250 citations to the published manual (as of early 2025), including papers describing further intervention work (e.g., Destin & Hernandez, 2021; Melguizo et al., 2021). For example, Florea and Negru-Subtirica (2025) replicate the positive effect *of Pathways-to-Success* on students' difficulty-as-importance and difficulty-as-impossibility beliefs; other teams document its impact on students' possible identities as engineers (Obery et al., 2023) and scientists (Kim et al., 2025).

Fit with What Works Clearinghouse recommendations

The United States Department of Education (DOE) set up the What Works Clearinghouse (WWC) to help identify intervention options and highlight the theoretically posited concepts empirically associated with better school outcomes. *Pathways-to-Success* addresses the only Recommendation in the latest WWC Practice Guide, *Preventing Dropout in Secondary Schools,* that the WWC views as backed by strong, not moderate or minimal, evidence (Rumberger et al., 2017). This Recommendation is to "engage students by offering curricula and programs that connect schoolwork with college and career success and improve students' capacity to manage challenges in and out of school." *Pathways-to-Success* does this through activities whose impact has been tested with a randomized control trial (RCT; Oyserman et al., 2006).

Prior rigorous sequential evaluations

Pathways-to-Success was initially manualized as an afterschool program, *School-to-Jobs,* implemented by undergraduate and graduate students in nine consecutive weekly sessions at the beginning of the school year (Oyserman et al., 2002). Compared to students attending other afterschool programs, those in *School-to-Jobs* significantly improved from baseline to end-of-year on the measured outcomes of school bonding and concern, frequency of being sent out of class, school absences, count of balanced school-focused possible selves, and plausible strategies (coded on a 0 to 4 scale from open-ended idiographic responses).

Next, *School-to-Jobs* was manualized as an in-school program and delivered in classrooms by community members (African American adults living in Detroit with undergraduate degrees in subjects other than education; for details on training and sample, Oyserman et al., 2006). This transition from after-school to in-school was accomplished by dividing session content into class periods for twice-weekly delivery across six weeks and testing effectiveness with a rigorous

National Institutes of Health-funded RCT (Oyserman et al., 2006). The RCT yielded meaningful and statistically significant positive impacts on end-of-8th-grade academic outcomes (decreased grade retention, increased attendance, better performance in core courses, and higher test scores) and possible identity content, structure, and strategy plausibility. Possible identity content and structure were coded from ideographic responses to open-ended probes and coding taken from Oyserman and colleagues (2004). Probes asked what the students expected they might become and wanted to avoid becoming in the coming year, if they were doing something now to work on these expectations and concerns, and if so, what they were doing. Skilled coders coded plausibility using the 6-point omnibus scale from Oyserman and colleagues (2004). Effects on academic outcomes persisted or grew larger by the end of 9th grade. Posited effects on difficulty-as-importance and difficulty-as-impossibility were not assessed as these scales had not yet been developed (initial experimental tests used priming only).

A third, DOE-funded follow-up showed that eighth-grade teachers could implement with fidelity, as assessed from a videotape of each session, and that fidelity mattered by yielding better academic outcomes as assessed by reduced risk of course failure and better grade point averages (Horowitz et al., 2018). A follow-up of this teacher-led version of *School-to-Jobs*, labeled *Pathways-to-Success,* with a larger sample, replicated this result (Oyserman et al., 2021). Moreover, fidelity positively affected IBM (as operationalized using the SVR scores published by Oyserman and O'Donnell, 2023, to code possible identities and strategy plausibility, and the difficulty-as-importance and difficulty-as-impossibility scale scores published by Elmore et al., 2016).

Open questions

We started with two open questions: do the positive effects of *Pathways-to-Success* on academic outcomes continue through a second follow-up year, and are they robust to ways of coding and analyzing the ideographic responses to open-ended possible identities and strategies probes? The staying power and robustness are the central practical questions, as detailed next, addressing robustness also advances research on the future self.

IBM theory predicts that in tandem, possible identities, plausible strategies, difficulty-as-importance, and difficulty-as-impossibility support the successful pursuit of academic goals. This theorized relationship among these aspects of IBM can be operationalized as a latent construct. Though not yet tried, a more accurate way would be to consider each aspect as part of a multi-faceted profile (a composite variable approach, see Bollen & Bauldry, 2011). Moreover, ideographic responses to open-ended probes provide insight into what is on students' minds rather than how much they agree or disagree with presented options (as is the case with closed-ended measures). Coding requires skilled coders. The advantage of skilled coders is offset by cost, error, and fatigue when they code at scale. The alternative is a machine-learning approach that yields a score reflecting the algorithm's certainty that the categorization fits requirements (an accuracy or prediction score). Prior approaches captured the content of possible identities and strategies either with prediction scores reflecting certainty of matching the human-coded material used to teach and test the algorithm (Horowitz et al., 2020) or SVR scores capturing the effect of possible identities and strategies scores on academic outcomes (O'Donnell & Oyserman, 2023). While useful, these approaches did not specifically operationalize plausibility in ways that parallel human coding.

Current study

We address the staying power (SRQ1) and robustness (SRQ2) open questions by adding 9th-grade data to the data reported by Oyserman and colleagues (2021). We provide our code in OSF: https://osf.io/h4xdj/?view_only=4eae0afeaf704db7a7ce783987745168

SRQ1 Staying Power: Does the indirect effect of fall 8th-grade *Pathways-to-Success* delivered with fidelity on spring 8th-grade academic outcomes via IBM carry over to spring of 9th-grade academic outcomes (Figure 1, top panel)?

SRQ2: Robustness of Effects: Are SRQ1 indirect effects robust to the method used to code and statistically represent IBM (Figure 1, middle and bottom panels)?

Materials and methods

Methods

Sample

Our initial sample, taken from Oyserman and colleagues (2021), included 8th-graders ($N = 1{,}142$, 51% girls, 87% free/reduced-price lunch) who participated in *Pathways-to-Success* in 40 classrooms in 10 high-poverty schools in the Chicago Public Schools (CPS) district (class size $M = 28.70$, $SD = 4.68$) during the academic years 2014–2015, 2015–2016, or 2016–2017. As detailed by Oyserman and

Figure 1. Predicted process model (top panel), IBM operationalized as a latent construct (middle panel) or as a composite profile construct (bottom panel).

colleagues (2021), students resembled the student population of urban districts (64% Hispanic or Latinx; 20% African American or Black; the remaining 16% White, Asian, Multiracial, or another race-ethnicity). Implementers ($N = 28$) were the 8th-grade teachers teaching one or more core subjects in the participating schools and trained by teachers who had implemented *Pathways-to-Success* with the highest fidelity in their cohort. The Oyserman and colleagues' (2021) analytic sample ($n = 1011$) dropped students lacking parental consent. The novel data for the current study are end-of-9th-grade administrative data for students still in the CPS system (final analytic $N = 875$, after excluding students lacking high school or IBM measurement data).

Data

Data collection procedures were detailed by Oyserman and colleagues (2021). Briefly, CPS provided administrative data as part of a data-sharing agreement with the American Institutes for Research. Students gave their feedback on the program approximately two days after their last session (late October or early November) and in late May or early June reported on their IBM (using the above-described structured open-ended responses for possible identities and strategies, six-item difficulty-as-importance scale, and six-item difficulty-as-impossibility scale). Before collecting data, the project obtained IRB approvals (Chicago Public Schools Research Review Board (CPS Project ID: 921) and USC (USC: UP 1400287).

Academic outcomes. The new data for this study were our predicted variables, 9th-grade grade point average (GPA), and risk of course failure. We followed Oyserman and colleagues' (2021) calculations of GPA and risk of course failure and used their data for 7th-grade (lagged control) and 8th-grade (proximal mediator) GPA and risk of course failure.

Fidelity scoring. We used the published fidelity scores from Oyserman and colleagues (2021). As an overview, teachers videotaped each session, and trained coders assessed each session recording using a structured set of rubrics. These rubrics and student feedback responses were compiled to create the five factors of fidelity (dosage, adherence, student responsiveness, quality of delivery, and student fidelity of receipt). The five factors were averaged to create an overall fidelity score. Individual scores for students nested in classrooms could be calculated because student responses were included in three factors (responsiveness, quality, and receipt).

School-focused possible identities and strategies. Large language models rely on co-occurrences of words and phrases to categorize content and provide a quantified certainty score that the sought-after content was recognized. As pertains to school-focused possible identities and strategies, the published 8th-grade results (Oyserman et al., 2021) used the coding approach provided by O'Donnell and Oyserman (2023) who used Word2Vec (Mikolov et al., 2013) to obtain support vector regression scores representing confidence that students' idiographic description of their possible identities and strategies are positively associated with subsequently better GPA. It applies sigmoid kernel regression to represent the mathematical relationship between how students describe their possible identities and linked strategies and their subsequently attained grades from school records. O'Donnell and Oyserman (2023) describe multiple checks on reliability for their Word2Vec approach; these include correlation with human-coded plausibility scores, and comparison of results from possible identity and strategy responses to results using other student-generated text.

In the current study, we developed an alternative, rule-based machine learning approach that more closely parallels how skilled human coders categorize possible identities as school-focused and assess the plausibility of their strategies to achieve their school-focused possible identities (Oyserman et al., 2004). We used the O'Donnell and Oyserman (2023) SVR approach to capture features of possible identities that predict academic outcomes. As a second latent construct, as detailed in the Appendix, we used large language models (LLMs) trained on human-coded data to create scores for the four aspects of possible identities and strategies that together instantiate plausibility. We used the LLM DistilBERT (Sanh et al., 2019) to get a vector representation of the content of each student's possible identities and strategies. This approach has two major advantages relative to Word2Vec. First, it produces a 768-dimensional vector that captures more subtle nuances of student writing content. Second, it considers the contextual meaning of words, a better fit to human coding of student responses. Indeed, our LLM-based functional score for the possible identities and strategies significantly correlated with GPA at $p < .001$ (7th-grade $r = .43$, 8th-grade $r = .42$, and 9th-grade $r = .32$, compared to the $r = 0.26$ for 8th-grade outcomes in Oyserman and colleagues' 2021 model).

Next, we used the LLM DeBERTa-v3-small (He et al., 2021) to create a machine-coded plausibility score comprising the four elements of plausibility used by skilled human coders in coding plausibility (described by Oyserman et al., 2004: the number of school-focused possible identities, having multiple (three or more) strategies to attain them, at least two of which are concrete, and at least one of which focuses on the social aspect of the school context (e.g., teachers, classmates). We computed continuous scores for each element to increase measurement sensitivity. While Oyserman and colleagues (2004) had skilled coders take an omnibus approach, we unpacked this into a set of criteria for operationalizing each plausibility element, and human coders coded responses using these criteria to train and test the LLM. The accuracy of the machine-coded score compared to the human-coded score ranged between 88.80% and 93.50% across the four elements, indicating high reliability.

Identity-based motivation. The published 8th-grade effects evaluation operationalized IBM using the outcome-based Word2Vec functional score described above and latent constructs representing difficulty-as-importance and difficulty-as-impossibility (Oyserman et al., 2021). Figure 1, middle panel, depicts this three-element latent IBM construct. We compare this three-element approach with the Bollen and Bauldry (2011) composite variable approach which allows us to include the relationships among our LLM-based plausibility variables, LLM-based possible identities functional score, and difficulty-as-importance and difficulty-as-impossibility constructs (see Appendix for the items) as shown in Figure 2, bottom panel.

Results

IBM measurement model: a profile analytic approach

We established a measurement model for IBM using our machine-coded plausibility score, LLM-based possible identities functional score, difficulty-as-importance, and difficulty-as-impossibility. We fitted a four-factor confirmatory factor analytic (CFA) model using the *lavaan* R package, as shown in Appendix Figure S1. We used full information maximum likelihood estimation to account for missing data (estimator set to "MLR"). This computed a scaled χ^2 statistic and sandwich estimates of standard errors that are robust to deviation from multivariate normality. We assessed model fit using goodness-of-fit indices including RMSEA, CFI, and SRMR (Hu & Bentler, 1998), finding an adequate fit, scaled χ^2 ($N = 875$, $df = 114$) $= 241.12$, $p < .001$, RMSEA $= .040$, 90% CI [.033, .047], CFI $= .975$, SRMR $= 0.032$. Having established fit, we used the Bartlett scoring method (e.g., Grice, 2001) to compute each student's 4-item IBM score estimate from the CFA model's factor scores. We used the method proposed by Lai and Hsiao (2022) to estimate and account for the measurement error of student-level IBM in our structural models.

Structural model

We ran two structural models, one for *SRQ1 Staying Power* and one for *SRQ2 Robustness of Effects*. The first used the composite IBM score to examine IBM's direct and indirect effects on

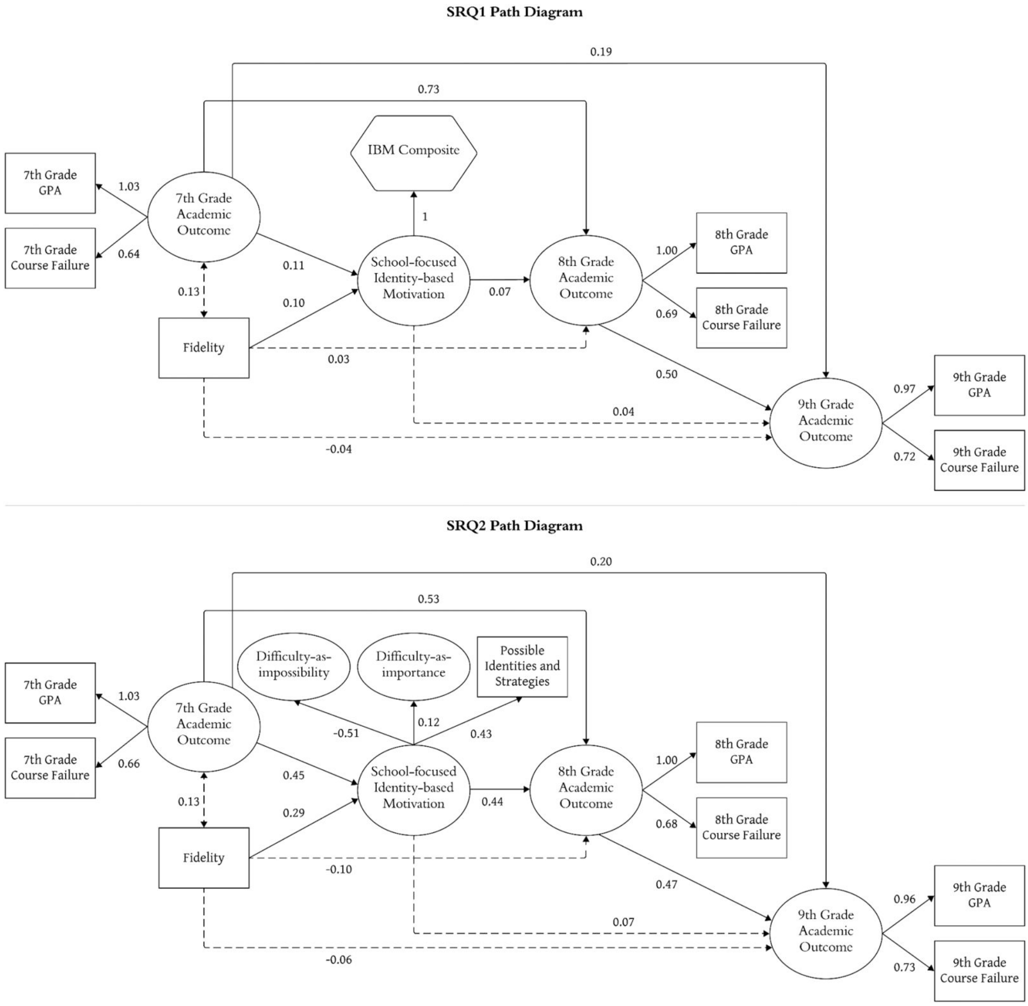

Figure 2. Path diagram for SRQ1 and SRQ2 with IBM operationalized as a composite construct (top panel) or as a latent construct (bottom panel). *Note.* Solid (dashed) paths indicate significant (non-significant) coefficients at $p < .05$. The numbers represent standardized coefficients. The top panel model has adequate fit, scaled χ^2 ($N = 875$, $df = 14$) $= 38.522$, $p < .001$, RMSEA $= .042$, 90% CI [.026, .058], CFI $= .984$, SRMR $= 0.032$. The bottom panel model has adequate fit, scaled χ^2 ($N = 875$, $df = 161$) $= 327.724$, $p < .001$, RMSEA $= .032$, 90% CI [.027, .037], CFI $= .972$, SRMR $= 0.049$.

9th-grade academic outcomes. The second substituted the composite IBM score for the latent IBM score to replicate the process model from Oyserman and colleagues (2021). We use *Mplus* 8.10 with maximum likelihood estimation, robust standard error, test statistics (i.e., the "ESTIMATOR = MLR" option), and cluster-robust standard errors ("TYPE = COMPLEX") to account for the nesting of students in classrooms in our data. We provide the related simple correlation matrices in the Appendix As Table A1 (latent factor model) and Table A2 (composite profile model).

SRQ1: Effects have staying power

As detailed in Figure 2, top panel, we found significant paths from 7th-grade academic trajectory to 8th-grade trajectory and lagged 9th-grade trajectory (controlling for the effects of 8th-grade). Students' 8th-grade academic outcomes positively predicted their 9th-grade academic outcomes

(β = .50, p < .001). Their 7th-grade academic outcomes also positively predicted 9th-grade outcomes above and beyond the effects of their 8th-grade outcomes (β = .19, p = .001). Paralleling Oyserman and colleagues (2021), IBM positively predicted 8th-grade outcomes (β = .07, p = .002). It had an indirect effect through 8th-grade outcomes with an estimated standardized indirect effect of β = .04, p < .05 (not a direct effect on 9th-grade outcomes (β = .04, p = .395). Paralleling the prior analyses (Oyserman et al., 2021), fidelity's effect on academic outcomes was also not direct (direct effect on 9th-grade academic outcomes (β = −0.00, p = .367); its effect was indirect *via* the path from IBM to 8th-grade academic outcomes. Prior academic trajectory and higher fidelity were associated with significantly higher IBM scores in the spring of 8th grade, and this was associated with significantly better 8th-grade academic outcomes that carried over to significantly better 9th-grade academic outcomes.

SRQ2: effects are robust to different ways of coding and operationalizing IBM

Our model provided an adequate fit to the data, whether we modeled IBM as a composite using LLMs to code two features of possible identities and strategies (Figure 2, top panel) or a latent construct (Figure 2, bottom panel). The basic pattern of significant effects is parallel between the two models, with prior academic trajectory and fidelity with which a student experienced *Pathways-to-Success* both associated with more IBM and with IBM and prior academic trajectories associated with better subsequent academic outcomes. Hence, we found evidence to support robustness. Results were consistent regardless of how the core constructs (idiographic possible identities and strategies) were coded or how possible identities and strategies and close-ended difficulty-as-importance and difficulty-as-impossibility responses are operationalized (statistically modeled) to represent IBM.

The two models appear to give different-sized coefficients for each path. However, the absolute values of the coefficients are not directly comparable across the two models because the models differently address measurement error. The relative sizes of coefficients within a model are also not comparable across models because the coefficients have different interpretations. Specifically, the coefficients in the latent IBM model represent values one would have obtained if one could measure the latent IBM construct without any error. While in the composite IBM model, the coefficients represent associations with the observed IBM components. Because there are three indicators, the latent IBM model makes a relatively big adjustment on the coefficients, resulting in more uncertainty in the coefficients, as reflected in the standard error (SE) estimates. Because of this, the Wald test statistics of the coefficients between the two models are more similar than they might seem: for example, in the composite IBM model, the standardized path coefficient from the IBM composite to 8th-grade academic outcomes was .07, SE = .023, z = 3.04, p = .002; in the latent IBM model, the standardized path coefficient from the latent IBM construct to 8th-grade academic outcomes was .44, SE = .093, z = 4.75, p < .001. Hence the composite IBM model implies that prior academic outcomes leave little room for other effects. In contrast, the latent IBM model implies that prior academic outcomes leave room for other effects and that IBM matters with approximately equal weight as prior academic outcomes. True effects are likely to include both possibilities; different ways of handling measurement errors are at least partially responsible for the apparent discrepancy between models.

Discussion

We started with an applied perspective on choosing school-based social and behavioral interventions: stakeholders want investments that last from the end of the intervention through the school year and onward through the next school year. They feel more confident if significance does not depend on a particular analytic technique or coding method. We examined *Pathways-to-Success*, a manualized intervention with published information on training and effects of the theorized active ingredient and of receiving the intervention with fidelity. We documented that these effects carry over from the end of one school year to the end of the next school year, from middle school to high school *via* the effects of fidelity on the postulated active ingredient, identity-based motivation. These effects are robust- they are not dependent on a single IBM operationalization or coding method. The mediation *via* IBM remains, even though the models differ in analytic choices- how they handle measurement error, the mechanics of coding possible identities and strategies, and the analytic choices in the construction of the IBM effect. The continuation and robustness of results to these variations increases confidence in *Pathways-to-Success*' effectiveness.

While our focus is practical, we also contribute to the theoretical literature by operationalizing IBM as a composite profile of four aspects: count of school-focused possible identities, scoring of school-focused strategies as a plausible roadmap, difficulty-as-importance, and difficulty-as-impossibility scores. Our operationalization is compatible with the theorized recursive connections among these aspects and yields parallel effects to a latent construct operationalization. Moreover, we provide the tools for districts to code and account for ideographic information on the number of school-focused possible identities students generate and the plausibility of their generated strategies as a roadmap to work toward these identities, as well as measures of difficulty-as-importance and difficulty-as-improvement. Taken together, our results increase confidence in the staying power of *Pathways-to-Success*, a relevant criterion for schools and districts considering *Pathways-to-Success* as a social behavioral intervention to support students' IBM, and in doing so, their academic outcomes.

Note

1. As detailed in Oyserman and colleagues (2012), Identity-based motivation theory uses the term identity rather than self to describe the content of the self, the term self-concept to describe an organizational structure of the self (e.g., independent self-concept, interdependent self-concept), and the term self to describe the object of reflection "me." In doing so, IBM theory attempts to reduce confusion as to whether the focus is on particular content (a current, past, or possible future identity) or on a more general feeling about the self (efficacy, optimism, esteem) that can occur when the same term is used for multiple meanings. For example, a possible self could mean a particular content, a general valence, or just the existence of an imagined future.

Acknowledgement

We thank Hitesh Manivannan for his work in developing the LLM, the involved schools, teachers, students, and especially our teacher trainers: Lahanna Addington, Cathy Cunningham-Yee, Sophia Panos, Joshua Silverman, and Jennifer Thompson.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The research reported in this paper was funded by the U.S. Department of Education, Institute of Education Sciences, grant number # R305A140281, to Oyserman and Sorensen.

ORCID

Daphna Oyserman http://orcid.org/0000-0002-2727-2167
Mark Lai http://orcid.org/0000-0002-9196-7406
Alysia Burbidge http://orcid.org/0000-0003-0113-1734

References

Allensworth, E., & Easton, J. (2007). *What matters for staying on-track and graduating in Chicago public high schools: A close look at course grades, failures, and attendance in the freshman year*. Research Report. Consortium on Chicago School Research. ERIC Number ED498350.

Bennett, D., Richardson, S., & MacKinnon, P. (2016). *Enacting strategies for graduate employability: How universities can best support students to develop generic skill Part A*. Australian Government, Office for Learning and Teaching, Department of Education and Training (http://www.olt.gov.au/project-how-universities-can-best-support-students-develop-generic-skills-enactingstrategies-gradua)

Bollen, K., & Bauldry, S. (2011). Three Cs in measurement models: Causal indicators, composite indicators, and covariates. *Psychological Methods*, *16*(3), 265–284. https://doi.org/10.1037/a0024448

Bradshaw, C., Zmuda, J., Kellam, S., & Ialongo, N. (2009). Longitudinal impact of two universal preventive interventions in first grade on educational outcomes in high school. *Journal of Educational Psychology*, *101*(4), 926–937. https://doi.org/10.1037/a0016586

Cook, P., Dodge, K., Farkas, G., Fryer, R., Jr., Guryan, J., Ludwig, J., Mayer, S., Pollack, H., & Steinberg, L. (2014). *The (surprising) efficacy of academic and behavioral intervention with disadvantaged youth: Results from a randomized experiment in Chicago*. (Working Paper No. 19862). National Bureau of Economic Research.

Dekker, I., & Meeter, M. (2022). Evidence-based education: Objections and future directions. *Frontiers in Education*, *7*, 941410. https://doi.org/10.3389/feduc.2022.941410

Destin, M., & Hernandez, I. (2021). The pathways intervention as a model to design broader systems of equitable student support. In G. Walton & A. Crum (Eds.), *Handbook of wise interventions: How social psychology can help people change* (pp. 148–165) The Guilford Press.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers) (pp. 4171-4186).

Durlak, J., Mahoney, J., & Boyle, A. (2022). What we know, and what we need to find out about universal, school-based social and emotional learning programs for children and adolescents: A review of meta-analyses and directions for future research. *Psychological Bulletin*, *148*(11–12), 765–782. https://doi.org/10.1037/bul0000383

Elmore, K., & Oyserman, D. (2012). If 'we'can succeed,'I'can too: Identity-based motivation and gender in the classroom. *Contemporary Educational Psychology*, *37*(3), 176–185. https://doi.org/10.1016/j.cedpsych.2011.05.003

Elmore, K., Oyserman, D., Smith, G., & Novin, S. (2016). When the going gets tough. *AERA Open*, *2*(3). https://doi.org/10.1177/2332858416664714

Efthymiadou, E., Gonida, E., & Kiosseoglou, G. (2025). Academic possible selves, motivational beliefs, and self-regulation among adolescents attending general and vocational schools: Does the type of school matter? *Behavioral Sciences (Basel, Switzerland)*, *15*(2), 158. https://doi.org/10.3390/bs15020158

Fisher, O., & Oyserman, D. (2017). Assessing interpretations of experienced ease and difficulty as motivational constructs. *Motivation Science*, *3*(2), 133–163. https://doi.org/10.1037/mot0000055

Florea, A., & Negru-Subtirica, O. (2025). Strengthening educational identity: Linking an identity-based motivation intervention to educational identity processes. *Identity*, *25*(2), 262–276. https://doi.org/10.1080/15283488.2024.2387600

Gage, N., Leite, W., Childs, K., & Kincaid, D. (2017). Average treatment effect of school-wide positive behavioral interventions and supports on school-level academic achievement in Florida. *Journal of Positive Behavior Interventions*, *19*(3), 158–167. https://doi.org/10.1177/109830071769355

Giner-Sorolla, R., Montoya, A. K., Reifman, A., Carpenter, T., Lewis, N. A., Aberson, C. L., Bostyn, D. H., Conrique, B. G., Ng, B. W., Schoemann, A. M., & Soderberg, C. (2024). Power to detect what? Considerations for planning and evaluating sample size. *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc*, *28*(3), 276–301. https://doi.org/10.1177/10888683241228328

Gresham, F. (2004). Current status and future directions of school-based behavioral interventions. *School Psychology Review*, *33*(3), 326–343. https://doi.org/10.1080/02796015.2004.12086252

Grice, J. (2001). A comparison of factor scores under conditions of factor obliquity. *Psychological Methods*, *6*(1), 67–83. https://doi.org/10.1037/1082-989x.6.1.67

He, P., Gao, J., & Chen, W. (2021). Debertav3: Improving deberta using electra-style pre-training with gradient-disentangled embedding sharing. ArXiv preprint. https://arxiv.org/abs/2111.09543

Horowitz, E., Oyserman, D., Dehghani, M., & Sorensen, N. (2020). Do you need a roadmap or can someone give you directions: When school-focused possible identities change so do academic trajectories. *Journal of Adolescence*, *79*(1), 26–38. https://doi.org/10.1016/j.adolescence.2019.12.013

Horowitz, E., Sorensen, N., Yoder, N., & Oyserman, D. (2018). Teachers can do it: Scalable identity-based motivation intervention in the classroom. *Contemporary Educational Psychology*, *54*, 12–28. https://doi.org/10.1016/j.cedpsych.2018.04.004

Hu, L., & Bentler, P. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*(4), 424–453. https://doi.org/10.1037/1082-989X.3.4.424

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, *55*(12), 1–38. https://doi.org/10.1145/3571730

Kase, C., Hoover, S., Boyd, G., West, K., Dubenitz, J., Trivedi, P., Peterson, H., & Stein, B. (2017). Educational outcomes associated with school behavioral health interventions: A review of the literature. *The Journal of School Health*, *87*(7), 554–562. https://doi.org/10.1111/josh.12524

Kim, M., DeRosia, N., Guldin, R., Lazaro, M., Husman, J., Madison, E., & Anderson, R. (2025). Perceptions and experiences of science among diverse adolescent youth during COVID-19. *Journal of Applied Developmental Psychology*, *97*, 101755. https://doi.org/10.1016/j.appdev.2025.101755

Lai, M., & Hsiao, Y.-Y. (2022). Two-stage path analysis with definition variables: An alternative framework to account for measurement error. *Psychological Methods*, *27*(4), 568–588. https://doi.org/10.1037/met0000410

Lee, J. (2024). The roles of Identity-Based Motivation and Perceived Instrumentality for probationary students' positive Self-Beliefs, Self-Regulation, and Performance. *Research in Higher Education*, *65*(4), 655–678. https://doi.org/10.1007/s11162-024-09800-y

Lewis, N., Jr.,., & Wai, J. (2021). Communicating what we know and what isn't so: Science communication in psychology. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *16*(6), 1242–1254. https://doi.org/10.1177/1745691620964062

Melguizo, T., Martorell, P., Swanson, E., Chi, W., Park, E., & Kezar, A. (2021). Expanding student success: The impact of a comprehensive college transition program on psychosocial outcomes. *Journal of Research on Educational Effectiveness*, *14*(4), 835–860. https://doi.org/10.1080/19345747.2021.1917029

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In C. Burges, L. Bottou, M. Welling, Z. Ghahramani, & K. Weinberger (Eds.), *Advances in neural information processing systems* (pp. 3111–3119). Curran Associates.

Obery, A., Lux, N., & Gannon, P. (2023). Designing elementary digital game-based engineering interventions for rural and Indigenous students. *International Journal of Engineering Education*, *39*(6), 1542–1555.

O'Donnell, S. C., & Oyserman, D. (2023). Apt and actionable possible identities matter: The case of academic outcomes. *Journal of Adolescence*, *95*(2), 354–371. https://doi.org/10.1002/jad.12118

Oyserman, D. (2007). Social identity and self-regulation. In A. Kruglanski & T. Higgins (Eds.) *Handbook of Social Psychology* (2nd ed., pp. 432–453). Guilford Press.

Oyserman, D. (2009). Identity-based motivation: Implications for action-readiness, procedural-readiness, and consumer behavior. *Journal of Consumer Psychology*, *19*(3), 250–260. https://doi.org/10.1016/j.jcps.2009.05.008

Oyserman, D. (2015). *Pathways to success through identity-based motivation*. Oxford University Press.

Oyserman, D. (2024). Identity-Based Motivation and the Motivational Consequences of Difficulty. *Social and Personality Psychology Compass*, *18*(12), e70028. https://doi.org/10.1111/spc3.70028

Oyserman, D., Elmore, K., & Smith, G. (2012). Self, self-concept and identity. Chapter 4 In M. Leary & J. Tangney (Eds.), *Handbook of self and identity* (2nd ed., pp. 69–104). Guilford Press.

Oyserman, D., & Horowitz, E. (2023). From possible selves and future selves to current action: An integrative review and identity-based motivation synthesis. *Advances in Motivation Science*, *10*, 73–147. ISSN 2215-0919 https://doi.org/10.1016/bs.adms.2022.11.003

Oyserman, D., Bybee, D., & Terry, K. (2006). Possible selves and academic outcomes: How and when possible selves impel action. *Journal of Personality and Social Psychology*, *91*(1), 188–204. https://doi.org/10.1037/0022-3514.91.1.188

Oyserman, D., Bybee, D., Terry, K., & Hart-Johnson, T. (2004). Possible selves as roadmaps. *Journal of Research in Personality*, *38*(2), 130–149. https://doi.org/10.1016/S0092-6566(03)00057-6

Oyserman, D., Johnson, E., & James, L. (2011). Seeing the destination but not the path: Effects of socioeconomic disadvantage on school-focused possible self content and linked behavioral strategies. *Self and Identity: The Journal of the International Society for Self and Identity*, *10*(4), 474–492. https://doi.org/10.1080/15298868.2010.487651

Oyserman, D., O'Donnell, S. C., Sorensen, N., & Wingert, K. M. (2021). Process matters: Teachers benefit their classrooms and students when they deliver an identity-based motivation intervention with fidelity. *Contemporary Educational Psychology*, *66*, 101993. https://doi.org/10.1016/j.cedpsych.2021.101993

Oyserman, D., Terry, K., & Bybee, D. (2002). A possible selves intervention to enhance school involvement. *Journal of Adolescence*, *25*(3), 313–326. https://doi.org/10.1006/jado.2002.0474

Premachandra, B., & Lewis, N. A. (2022). Do we report the information that is necessary to give psychology away? A scoping review of the psychological intervention literature 2000–2018. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *17*(1), 226–238. https://doi.org/10.1177/1745691620974774

Rawte, V., Chakraborty, S., Pathak, A., Sarkar, A., Tonmoy, S. M., Chadha, A., … Das, A. (2023). The troubling emergence of hallucination in large language models–an extensive definition, quantification, and prescriptive remediations. ArXiv. https://arxiv.org/abs/2310.04988

Rumberger, R., Addis, H., Allensworth, E., Balfanz, R., Bruch, J., Dillon, E., Duardo, D., Dynarski, M., Furgeson, J., Jayanthi, M., Newman-Gonchar, R., Place, K., & Tuttle, C. (2017). *Preventing dropout in secondary schools*. (NCEE 2017-4028). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. https://whatworks.ed.gov

Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2019). DistilBERT, a distilled version of BERT: Smaller, faster, cheaper and lighter. *ArXiv*. https://arxiv.org/abs/1910.01108

Yan, V. X., Oyserman, D., Kiper, G., & Atari, M. (2024). Difficulty-as-improvement: The courage to keep going in the face of life's difficulties. *Personality & Social Psychology Bulletin*, *50*(7), 1006–1022. https://doi.org/10.1177/01461672231153680

Zhang, Y., Cook, C., Azad, G., Larson, M., Merle, J., Thayer, J., Pauls, A., & Lyon, A. (2023). A pre-implementation enhancement strategy to increase the yield of training and consultation for school-based behavioral preventive practices: A triple-blind randomized controlled trial. *Prevention Science: The Official Journal of the Society for Prevention Research*, *24*(3), 552–566. https://doi.org/10.1007/s11121-022-01464-3

Appendix

Large language model approaches

In coding ideographic responses to open-ended possible identity and strategy probes, we paired two large language model approaches to capture the complexity of open-ended possible identity and strategy responses. First, a **functional possible identity** approach based on students' possible identities and strategies to capture the content and structure of possible identities that predict students' grades. Second, a **rule-based plausibility** approach decomposed the operationalization of possible identity plausibility into four simple rules that undergrads could follow. Next, we detail the steps involved in each approach.

The "functional possible identity" approach

We obtained transformer word embeddings for possible identities and strategies using DistilBERT. The output from DistilBERT was 768 numerical values representing what students wrote for their possible identities and strategies. The 768 numerical values were used as the input values in a support vector regression analysis to predict grades. We trained the model to predict grades using k-fold validation with 10 folds. The model is trained on k-1 of these folds (in this case, 9 folds) and tested on the remaining fold. This process is repeated k times, with each fold serving as the test set exactly once. This allowed us to obtain predicted values for our entire dataset without rendering part of the dataset unusable as a training dataset.

We used these LLM-based functional scores in our composite profile construct to represent the functional relationship between students' possible identities and strategies on the one hand and their grades on the other. To do so, we employed an iterative modeling approach using SVR with a sigmoid kernel, evaluating the model using stratified 5-fold cross-validation, optimizing for Root Mean Square Error (RMSE) performance. The final hyperparameters selected were a gamma of 0.0036333, a regularization parameter Cost of 3.7, and an epsilon of 0.00. The embeddings derived from DistilBERT provided contextual information to capture the underlying patterns in students' possible identities and strategies. This allowed the model to balance accuracy and generalization across contexts. Figure A1 shows that the possible identities content functional score and the possible identities plausibility construct are moderately correlated and not redundant.

The rule-based plausibility approach

Plausibility has previously been coded as an omnibus score requiring a high level of training. We expanded the previously published coding dictionary so that we could train three undergraduates to code the four questions embedded in the omnibus score: Were there school-focused possible

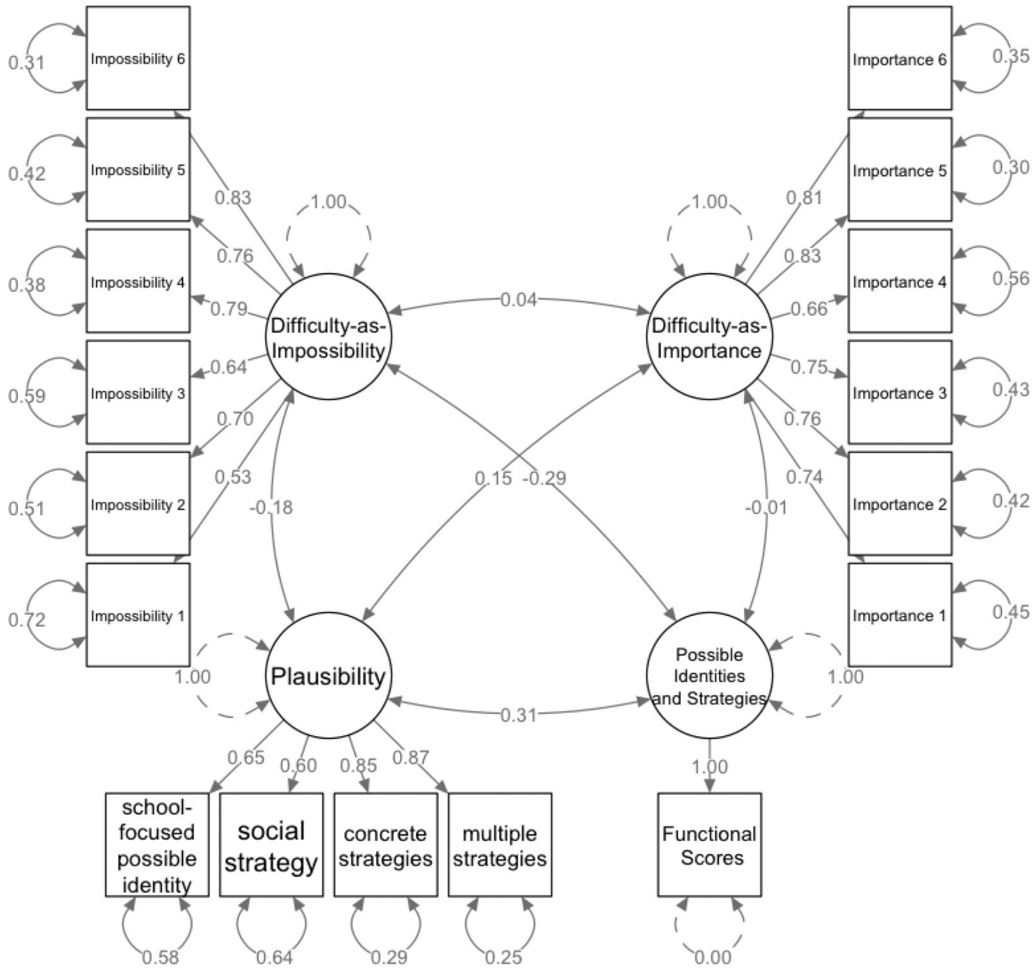

Figure A1. Identity-based motivation as a four-factor relationship: Confirmatory Factor Analysis.
Note: The numbers represent standardized coefficients. The four-factor model has adequate fit, RMSEA = .040, 90% CI [.033, .047], CFI = .975, SRMR = 0.032.

identities (if yes, then count; if no, then score 0 and stop)? Were there at least three strategies to attain the school-focused possible identities? (yes, no) Were at least two of these strategies specific and concrete? (yes, no) Did at least one of these strategies consider social context? (yes, no). We used the coding dictionary to train undergraduates to attain sufficient inter-rater reliability (IRR) with each other before they could begin independent coding. We aimed for a minimum of 80% agreement in their categorizations of each plausibility element. Coders discussed discrepant codes to agreement under the supervision of a member of the author team to ensure comprehension and accuracy. Undergraduates coded the responses of $n = 1844$ students using the same training data sets as used in Oyserman and colleagues (2021) and baseline data from a third data set (IES Goal 2 Grant #R305A140281, I3 Grant #U411C150011, EIR (Pilot) Grant: S411B200027). As a result of the IRR process, about 20% of the dataset was double-coded.

We used these human-generated codes to train our machine algorithms for the rules and checked whether these four codes were best considered separate factors or a single score.

The first rule was like a funnel: students who scored 0 on rule 1 received 0 as codes on all other rules. We trained our models by building on advances in large language models that outperform traditional natural language processing models (e.g., recurrent or convolutional neural networks Devlin et al., 2019). We trained various LLMs including BERT, roBerta, T5, Electra to complete the task, finding that DeBERTa-v3-small had the most consistent results for the dataset. We used RayTune to conduct hyperparameter tuning and 5-fold validation for the final machine coding. The machine algorithm provides a binary output of score 0 or 1 and a continuous value approximately between −4 and +4 representing its confidence in making the scoring classification.

The more positive the score is, the more confident the algorithm is that the rule is scored 1 and the more negative the score the more confident the algorithm is that the rule is scored 0. This confidence score provides an advantage of creating more granularity in the coding, hence we used the confidence score for each aspect of plausibility in our analysis.

Using responses from students with complete data, the machine algorithm was 93.50% accurate compared to human code. Accuracy was 92.00% for multiple strategies to work on school-focused possible identities, 88.80% for whether these strategies were specific, concrete, and actionable, and 91.7% for whether at least one of these strategies dealt with the social context entailed in making progress on school-focused possible identities. Accounting for trickle-down inaccuracy in coding, we estimated the practical accuracy of coding future data in the range of 85% to 91% across the four codes.

Creating the difficulty-as-importance and difficulty-as-impossibility latent constructs

Students read and rated (1 = *strongly disagree* to 5 = *strongly agree*) these statements from Elmore and colleagues (2016).

Difficulty-as-importance

1. If I'm working on a task that feels difficult, it means that the task is important.
2. A sign that a task is important to me is how difficult it feels while working on it. If it feels difficult, it's important.
3. Struggling to complete a task reminds me that the task is important.
4. If a task is difficult, it is probably important for me to do well at it.
5. Tasks that feel difficult are important tasks.
6. If a task is difficult, it means that it's important for me.

Difficulty-as-impossibility

1. If I feel stuck on a task, it's a sign that my effort is better spent elsewhere.
2. If working on a task feels very difficult, that type of task may not be possible for me.
3. If a task feels too difficult, I should move on to something else.
4. When working on a task feels hard, that feeling means it's not for me.
5. Finding a task really difficult tells me that I can't complete that task.
6. If a task feels really difficult, it may not be possible for me.

Confirmatory factor analysis of the 4-factor identity-based motivation profile

Table A1. Correlations Between Variables in the IBM Latent Construct Model.

	Correlations						
B	1	2	3	4	5	6	7
1. 7th Grade Academic Outcome	–	.738	.574	−0.244	−0.019	.125	.234
2. 8th Grade Academic Outcome	<.001	–	.657	−0.357	.034	.122	.281
3. 9th Grade Academic Outcome	<.001	<.001	–	−0.235	−0.013	.049	.213
4. Difficulty-as-Impossibility	<.001	<.001	<.001	–	.036	−0.153	−0.233
5. Difficulty-as-Importance	.665	.383	.002	.511	–	.186	.094
6. Fidelity	.079	.074	.228	.001	<.001	–	.147
7. Possible Identity Score (Word2Vec + SVR)	<.001	<.001	<.001	<.001	.004	<.001	–

Note. The upper triangle shows correlations between latent and observed variables included in the IBM latent construct model. The lower triangle shows *p*-values.

Table A2. Correlations between variables in the IBM composite path model.

	Correlations							
Measure	1	2	3	4	5	6	7	8
1. 7th Grade Academic Outcome	–	.738	.574	−0.244	−0.019	.112	.125	.419
2. 8th Grade Academic Outcome	<.001	–	.657	−0.357	.034	.245	.122	.440
3. 9th Grade Academic Outcome	<.001	<.001	–	−0.235	−0.013	.222	.049	.327
4. Difficulty-as-Impossibility	<.001	<.001	<.001	–	.036	−0.193	−0.153	−0.290
5. Difficulty-as-Importance	.665	.383	.002	.511	–	.155	.186	−0.003
6. LLM-Coded Plausibility Score	.012	<.001	<.001	<.001	<.001	–	.114	.312
7. Fidelity	.079	.074	.228	.001	<.001	.052	–	.102
8. Possible Identities Score (LLM + SVR)	<.001	<.001	<.001	<.001	.922	<.001	.062	–

Note. The upper triangle shows correlations between latent and observed variables included in the IBM composite path model. The lower triangle shows *p*-values.