Prevention Science
https://doi.org/10.1007/s11121-020-01097-4

Effect of Daily Teacher Feedback on Subsequent Motivation and Mental Health Outcomes in Fifth Grade Students: a Person-Centered Analysis

Aaron M. Thompson[1] · Wolfgang Wiedermann[1] · Keith C. Herman[1] · Wendy M. Reinke[1]

Abstract

Prevention scientists recognize that implementing effective prevention practices and programs responsive to the needs of individuals but based solely upon the findings from variable-centered methods presents several limitations due to numerous risk factors, pathways, and unobserved influences. One such understudied influence that is masked by variable-centered methods, motivation, is a person-level characteristic that influences treatment outcomes. The purpose of this paper is to demonstrate the use of an alternative person-centered approach, group iterative multiple model estimation (GIMME), to model change over time that focuses on the interdependence of daily student motivation levels and teacher feedback and their relations to student outcomes over time. Specifically, we used GIMME to model person level responses to negative teacher feedback regarding students' next day motivational ratings using data from 58 5th grade students participating in a study of the impact of the self-monitoring and regulation training strategy (SMARTS). Results identified a set of SMARTS students whose daily readiness aligned with high rates of self and teacher agreement regarding ongoing performance ratings. However, results identified a group of students whose daily motivation and readiness for change was adversely impacted by negative teacher feedback the day before. For these students, they were more likely than their peers to experience high levels of depression and internalization scores. Motivationally oriented practice suggestions for providing feedback to students who may be sensitive to this type of feedback and research implications of these findings are discussed.

Keywords Motivation · Person-centered · Intervention · Externalization · Internalization

Prevention scientists have long recognized implementing effective prevention practices, and programs responsive to the needs of users but based solely upon the findings from variable-centered methods presents several limitations. First, due to numerous risk factors and pathways that interact with individual characteristics and contextual influences to condition a person's response to a treatment program, variable-centered methods mask our capacity to detect nuanced differences in individual responses to any particular prevention program. This is a problem because no single program adequately targets all possible mechanisms in an individually tailored manner. Second, the effect of motivation on treatment outcomes—a person-level characteristic—is widely understudied, particularly from a person-centered approach (Bergman and Magnusson 1997; Wiedermann et al. 2016). Even if the effect of individual motivation on outcomes was examined, the true effect of motivation on treatment uptake would likely remain masked in variable-centered analyses examining moderators. The purpose of this paper is to demonstrate the use of an alternative person-centered approach to model change over time that focuses on the interdependence of daily student motivation levels and teacher feedback and their relations to student outcomes over time.

With the dissemination of successful prevention programs, there has also been a growth in the use of prevention science technologies in the daily practices of professionals working with youth in school settings. For example, there is a growing use of individual risk screening and assessment tools and procedures to identify youth with varying levels of risk (Dowdy et al. 2010). There is also increasing wide-spread adoption of tiered prevention and intervention frameworks which

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11121-020-01097-4) contains supplementary material, which is available to authorized users.

✉ Aaron M. Thompson
thompsonaa@missouri.edu

[1] University of Missouri, 718 Clark Hall, Columbia, MO 65211, USA

Published online: 13 February 2020

integrate a range of programs and practices in an effort to combat an array of risk factors experienced at varying degrees by individuals in a population of youth (e.g., Positive Behavior Interventions and Supports; Flannery et al. 2014). Beyond prevention science technologies, there is also an increasing awareness of the influence of varying levels of individual motivation and how that predicts the success of prevention programs aiming to reduce aggression and peer victimization (Cunningham et al. 2012) or improve academic motivation (Wigfield and Cambria 2012).

Motivation or readiness to change is a common topic in studies examining precursors for change in the adult literature. Miller and Rollnick (1991) defined motivation as "the probability that a person will enter into, continue, and adhere to a specific change strategy." Motivation in this view is measured by asking the individual to report their readiness to engage in a behavior change action plan. This conceptualization of motivation helps explain heterogeneity of individual behavioral enactments as well as response to treatments. In line with this definition, López-Viets et al. (2002) specified six key characteristics of motivation: (1) it is modifiable, (2) the relation between motivation and action is characterized by a probability, (3) motivation is influenced by the context in which it is discussed, (4) motivation is specific to each intended course of action and thus may vary based on the specific behavior being considered, (5) motivation is intrinsic and extrinsic, and (6) intrinsic motivation is more likely to increase when it is elicited from the individual rather than others.

Motivational interviewing (MI; Miller and Rollnick 2012) is a clinical technique based on this definition and these principles of motivation and is a widely applied and evidence-based approach to absolving ambivalence and evoking readiness to change. Helpers use statements conditioned to elicit client change talk—the primary mechanism leading to behavior change in MI (McNamara 2009; Miller and Rose 2009). One statement MI helpers use is to provide personalized feedback to a client about ongoing performance related to desirable outcomes. The purpose of personalized feedback is to (a) verify whether a gap exists between current and expected performance levels, (b) ascertain the magnitude of the gap, (c) elaborate strategies to reduce the gap, and (d) assist with monitoring ongoing performance (Shute 2008). As one example applied in an educational context, Reinke and colleagues developed the MI-based classroom check-up (CCU) to give teachers specific feedback about specific domains of effective classroom management (e.g., rate of praise versus reprimands, rate of opportunities to respond; Reinke et al. 2011). Consultants present feedback to evoke change talk about what the teacher wants to be different. Several studies indicate that CCU improves use of best teaching practices (Reinke et al. 2008; Reinke et al. 2011).

Although MI informed that approaches have been less commonly used in interventions for youth, similar principles apply (Herman et al. 2013). Many educational interventions employ personalized feedback to monitor and foster student progress in acquiring math and reading skills (Deno et al. 2009; Fuchs et al. 2007). Reading interventions commonly use daily reading probes of performance skills like number of words-read-per-minute, and this data is charted and presented to the student (Fuchs and Fuchs 1999). Check in/check out (Crone et al. 2010) is a widely used tier 2 intervention that provides students with teacher daily ratings of student behaviors that are targeted for improvement. Thus, formative feedbacks are staples of educational interventions targeting both academic and social-behavioral outcomes with much literature supporting their effectiveness (Shute 2008; Reinke et al. 2008; Reinke et al. 2007).

An untested assumption of performance feedback is that all youth benefit from this type of feedback equally. Additionally, although performance feedback in educational contexts is consistent with MI, most existing educational applications of feedback are not informed by MI but are often driven by a behavioral orientation that hypothesizes that feedback increases awareness which by itself leads to be behavior change. Because feedback in school settings is nearly always provided by adults to students, it may be that the way adults present the feedback influences how it is received or that student characteristics influence the uptake of feedback. For instance, MI theory suggests that motivation is influenced by the interpersonal context in which it is discussed. Thus, although personalized feedback can evoke higher levels of motivation when delivered in a positive and supportive context, the opposite effect can occur when feedback is delivered in a judgmental and non-supportive context the probability of behavior change.

An additional factor that may influence youth response to feedback is their pre-existing self perceptions of their competence and skills. MI theory predicts that self perceptions of competence may influence an individual's response to feedback as well as their motivation or readiness to change a behavior. Evidence suggests that some youth with challenging behaviors, including both externalizing and internalizing behaviors (Kaiser and Ramninsky 2009), hold inaccurate perceptions of performance (see Hoza et al. 2004; Owens et al. 2007). For instance, a significant body of research has found that youth with challenging behaviors hold what is known as a *positive illusory bias* (Hoza et al. 2004) where they overestimate their competence compared with the ratings by others. Unlike modest illusory biases observed in the general population, the bias exhibited by youth with challenging behaviors is much more extreme and does not appear to foster or promote motivation to change. Although many explanations for the origins of this illusory bias have been offered (immaturity, skill deficit), evidence suggests that the bias serves a self-protective purpose where youth with challenging behaviors overestimate ability to protect their social and self image (Owens et al. 2007).

Given these biases and purported mechanisms, it is important for feedback interventions for youth with challenging behavior problems to examine variation in treatment responses. Feedback which provides youth with information about discrepancies or agreement between their ratings of performance compared with others' ratings could evoke positive or negative effects on youth motivation and subsequent performance dependent upon individual characteristics. Performance feedback is intended to narrow the discrepancy but could serve to undermine subsequent performance and motivation in youth threatened by the feedback's unmasking of their illusory bias self-image protection. One intervention, the *self-monitoring and regulation training strategy* (*SMARTS*), is guided by the principles of motivational interviewing and provides performance and discrepancy feedback to students with challenging behaviors.

SMARTS

The SMARTS (formerly called STARS; Thompson 2014; Thompson and Webber 2010) is a manualized tier 2 support intervention for youth with challenging behaviors that includes many of the elements of performance feedback. SMARTS includes features of existing selective programs that use formative adult feedback to drive behavioral change (i.e., *Check*, *Connect*, and *Expect*; *The Behavior Education Program*; *check in/check out*), but SMARTS also attempts to promote autonomy supportive opportunities for students to self monitor their own goal performance and compare their own perspectives with that of teachers. SMARTS structures direct instruction in key social-emotional skills to shape student goal formation, self monitoring, and processing of discrepancy feedback. Following student training, SMARTS structures daily opportunities for students to monitor social-emotional goals, compare self and teacher performance feedback, and consider strategies to reduce discrepancies between self and teacher perspectives. Self monitoring is a widely used motivational support strategy. Theoretically, self-monitoring is an autonomy support activity which is a key element in self-determination theory (Deci and Ryan 2011). When educators use autonomy support strategies like self monitoring, they promote student choice, directly involve students, increase opportunities for students to practice, and structure more opportunities for educators to clarify expectations for classroom behavior (Wigfield et al. 2008; Wentzel 2002) which has been shown to directly increase motivation to perform the expected behaviors via the concepts of engagement and reactivity (Pintrich & Schunk, 2002). Reactivity is a metacognitive principle referring to behavioral change occurring as a function of observing one's own behavior. Observed in studies of self monitoring of learning, students who self monitor math performance gain increased awareness of the number of problems they have correct and engage in strategies that directly improve performance (Bandura 2005; Cleary and Zimmerman 2004). The reactivity principle has also been observed in self-monitoring as an intervention in weight loss and substance abuse studies (Boutelle and Kirschenbaum 2012; Butryn et al. 2012; Sinadinovic et al. 2010). Targeted interventions like SMARTS that focus on tapping motivational strategies and utilize performance feedback to encourage students to learn, adopt, and utilize adaptive behaviors at school are individualized approaches to behavior support. Individualized approaches in motivation are useful in that it permits school personnel to tailor goals and feedback to meet student needs. However, these individualized interventions also present opportunities for alternative methods for analyzing change.

The purpose of our paper is to discuss group iterative multiple model estimation (GIMME; Gates and Molenaar 2012; Beltz and Gates 2017) as a subject-specific approach to evaluate dynamic relations in intensive longitudinal data. GIMME uses structural equation models (SEMs) that account for sequential dependence observed in time series data to identify population-, shared-, and individual-level relations among variables. The present study uses GIMME to evaluate the dynamic relation between daily teacher feedback on student behavioral goal performance and student motivation and its impact on student mental health outcomes. We hypothesized that multiple causal classification patterns would emerge from the analyses including subgroups of students for whom feedback affected subsequent student motivation and others for whom motivation affected subsequent feedback. Additionally, we examined the classification patterns' relation to student outcomes including student-reported externalizing and internalizing symptoms. We hypothesized that negative classification patterns, that is, those marked by a decline in student motivation or an increase in negative feedback, would predict worse student emotional outcomes. Finally, we determined the relations between baseline student-teacher relations and classification patterns and hypothesized that negative baseline student-teacher relations would predict negative classification patterns.

Method

The sample used in the present study was collected as part of an ongoing randomized control study to evaluate SMARTS. Two hundred ninety-three students were recruited for participation in the study using a dual-gated screening and assessment system. The first gate relied on the behavioral and emotional screening system (BESS; Kamphaus and Reynolds

2007) to screen 100% ($N = 2059$) of 5th grade students. Students identified as high risk on the BESS then were assessed at the second gate using the behavioral assessment system for children (Reynolds et al. 2011). Gate 2 inclusion criteria required students to exhibit a T score at or above 60 on the externalizing, internalizing, or school problems scales. As such, the present study included 154 students with elevated risk across three cohorts (2016: $n = 49$ students, 2017: $n = 59$ students, 2018: $n = 46$ students, see online supplement) who were randomly assigned to the SMARTS intervention only (i.e., control condition students are not included). Twenty-three students (14.9%) were excluded due to not meeting the gate 2 T score requirement.

Of the remaining 131 students (11.1 years of age; $SD = 0.4$) in the sample, 73.3% qualified for free or reduced lunch and self-identified as White (43.5%) Black (41.2%), Biracial (8.4%), Hispanic (4.6%), or Asian (2.3%). The sample included more male than female students (71.8%) and 25.8% received a special education service. Students showed average internalizing and externalizing T scores of 64.8 ($SD = 15.0$) and 72.6 ($SD = 14.9$), respectively (further descriptive statistics are given in the online supplement). Following Wright et al. (2019) and Lane et al. (2019), 72 students (55.0%) who had more than 40% missing values in daily self monitoring and teacher feedback on individual goals were discarded from this study. A cutoff of > 40% was a reasonable trade-off between convergence rates of individual time series models and retaining as many students as possible for subsequent (variable oriented) analyses of students' mental health outcomes. In addition, one student reported no variability in daily motivation and was excluded, leaving this analysis to be conducted with a sample of $n = 58$ students. Students in the analysis sample were less likely to qualify for free or reduced lunch (63.8%) than students in the subsample with > 40% incomplete data (80.8%, $p = 0.047$) and tended to show lower baseline depression (analysis sample $M = 68.0$, $SD = 14.7$; excluded subsample $M = 72.9$, $SD = 15.7$; $p = 0.071$) and follow-up externalization T scores (analysis sample $M = 66.1$, $SD = 14.3$; excluded subsample $M = 70.7$, $SD = 15.5$; $p = 0.086$). No significant differences between the analytical sample and excluded subsample were observed for age, race, gender, special education service, cohort, baseline externalizing and internalizing T scores, and follow-up internalizing and depression T scores. Table 1 (third column) gives the descriptive statistics for the sample.

Intervention Procedures

Delivery of SMARTS is organized in three phases: *training*, *self monitoring and feedback*, and *processing*. SMARTS training is delivered by student support personnel (school psychologists counselors, social workers) in a small group sessions lasting about 35–40 min each using ten scripted lessons in: (1) pre-group meeting, (2) assessing and defining problems, (3) generating and considering solutions, (4) writing measureable goals, (5) observing and recording progress, (6) using data to evaluate progress, (7) taking the perspective of others, (8) reframing mistakes as part of learning, (9) managing internal responses to problems, and (10) managing external responses to problems. Following training, students begin daily self monitoring where both students and teachers monitor students' performance each hour of the day on their individualized goals. Each day, however, before students and teachers rate ongoing goal performance, students respond to three motivational prompts scaled 0—"*not at all*" to 10—"*very ready*," measuring (a) how rested they feel, (b) how positive their mood is, and (c) how ready they are to accomplish their goal. Next, students and teachers select one of three response options ("*yes*," "*sometimes*," or "*no*") reflecting the student's goal performance during the prior 1-h interval. Teachers' recording of student behavior performance is similar to other tier 2 approaches (i.e., CICO). Unique to SMARTS and not CICO, SMARTS students are trained in goal development and self-monitoring skills, they use a website to chart and record performance, and can view graphs and percentages of their self and teacher ratings of their progress. Lastly, phase 3 consists of a weekly processing meeting between a school counselor and SMARTS students where they discuss in greater detail the daily performance feedback collected over the week. Using the web-based dashboard percentages and graphics, SMARTS students compare self and teacher observational data by responding to several motivational interviewing prompts encouraging students to compare present data with (a) their current goal, (b) prior goals and performance data, and (c) concurrent teacher data. During processing, students also consider behaviors that contributed to the discrepancies. Using the data and behavioral information, students then revise their goals each week. The revised goal is then entered into the website using observable and measureable language, and the student and teacher performance monitoring and processing are repeated iteratively. All three phases were implemented similarly and at the same time during the school year across each of the three cohorts. Median time series lengths were 32 ($Q_1 = 30$, $Q_3 = 32$) out of 32 possible days in 2016, 39.5 ($Q_1 = 33.8$, $Q_3 = 41$) out of 41 possible days in 2017, and 38 ($Q_1 = Q_3 = 38$) out of 42 possible days in 2018.

Measures

Following gate one screening, consent/assent, and gate two pretests, students were randomly assigned to SMARTS or a control. Posttest surveys were completed following SMARTS training and 8 weeks of self and teacher monitoring. In the analysis, we used daily student readiness ratings and daily rate

Table 1 Demographic information and descriptive statistics of the analysis sample

Variable		NTF → SM ($n = 12$)		SM → NTF or no relation ($n = 46$)		Total($n = 58$)	
Female	n (%)	3	(25.0)	11	(23.9)	14	(24.1)
Black	n (%)	5	(41.7)	17	(37.0)	22	(37.9)
FRL	n (%)	10	(83.3)	27	(58.7)	37	(63.8)
SPED	n (%)	4	(33.3)	11	(23.9)	15	(25.9)
Cohort	n (%)						
2016		6	(50.0)	13	(28.3)	19	(32.8)
2017		5	(41.7)	13	(28.3)	18	(31.0)
2018		1	(8.3)	20	(43.5)	21	(36.2)
Age (in years)	M (SD)	11.2	(0.4)	11.1	(0.5)	11.1	(0.5)
NTF	M (SD)	0.61	(0.44)	0.47	(0.42)	0.50	(0.43)
Student motivation	M (SD)	8.46	(2.91)	8.31	(2.37)	8.34	(2.49)
Externalization (pre)	M (SD)	70.3	(10.1)	71.5	(13.1)	71.2	(12.5)
Internalization (pre)	M (SD)	67.4	(13.3)	61.1	(13.9)	62.4	(13.9)
Depression (pre)	M (SD)	69.3	(14.4)	67.7	(15.0)	68.0	(14.7)
Externalization (post)	M (SD)	69.0	(15.0)	65.4	(14.2)	66.1	(14.3)
Internalization (post)	M (SD)	69.0	(10.1)	57.9	(13.8)	60.2	(13.8)
Depression (post)	M (SD)	74.3	(17.2)	63.0	(15.4)	65.3	(16.4)

M, mean; *SD*, standard deviation; *pre*, pre-treatment measure; *post*, post-treatment measure; *FRL*, free/reduced lunch; *SPED*, special education service; *NTF*, proportion of daily negative teacher feedback

of negative teacher feedback as predictors of social-emotional outcomes and student rated relations with their teacher.

Student Daily Readiness

Student participants rated daily readiness for change using three MI rulers patterned after Miller and Rollnick (1991) asking "*How positive do you feel today*?", "*How well rested do you feel today*?", *and* "*How ready are you to follow your goal today*?" with responses ranging from 0 (low readiness) to 10 (high readiness). Similar to studies using MI rulers to score daily readiness for change, the three items here were combined into a single composite ($\alpha = 0.86$) representing daily readiness for change. Readiness for change rulers are brief measures of a person's motivation for change and are highly correlated with multi-item readiness questionnaires measuring readiness for change in studies of alcohol and substance use treatment (0.77; LaBrie et al. 2005), smoking cessation (0.87; Boudreaux et al. 2012), and safe sex prevention practices (0.77; LaBrie et al. 2005).

Negative Teacher Feedback

Negative teacher feedback was measured as the daily relative frequency of cases where a positive performance rating was observed by the student but a negative performance rating was entered by the teacher (i.e., the student and teacher disagreed on whether the student achieved the goal). Frequencies ranged from 0 to 1 with higher scores indicating more greater rates of daily negative teacher feedback.

Social-Emotional Health Outcomes

Student social-emotional health was measured using the behavior assessment system for children 3—teacher rating system (BASC; Reynolds et al. 2011). Specifically, we relied on BASC subscales that measured *internalization problems* ($\alpha =$ 0.85– 0.89) and *externalization problems* ($\alpha = 0.83$ –0.86). We also used the BASC 3 self-report of personality to ask youth about their *depressive* symptoms ($\alpha = 0.93$).

Student-Teacher Relations

Students responded to the teachers who care scale of the elementary school success profile (ESSP-Student; 5 items, $\alpha =$ 0.72; Bowen 2010) measuring perceptions of whether a teacher listened, praised, provided help, and got along with the student.

Analysis

We used GIMME to evaluate longitudinal patterns of associations of negative teacher feedback and student motivation. GIMME is a data-driven approach that allows to distinguish population, shared, and individual effects from intensive longitudinal data. Population effects are conceptualized as effects that occur for the majority of individuals comprising the sample (effects are allowed to vary across individuals). Shared effects are those that occur for subgroups of the sample, and individual effects reflect processes that are unique to the given individual.

The GIMME algorithm is based on a unified structural equation modeling (uSEM; Kim et al. 2007; Lane et al. 2019; Molenaar 2019) framework linking standard SEMs and vector autoregressive (VAR) models (Lütkepohl 2005) and provides lagged and concurrent effects of negative teacher feedback and student motivation. Lagged effects refer to cases where measures of one construct (e.g., teacher feedback) at time point t predicts another construct (student motivation) at the subsequent time point $t+1$. Concurrent effects describe direct effects of one construct on the other the same day (e.g., negative teacher feedback affects student motivation at time t) and occur when the underlying data-generating mechanism changes faster than the rate of data collection (Granger 1969). In addition, autoregressive effects are estimated which quantify how well a variable predicts itself over time. Such autoregressive effects can be conceptualized as measures of stability. Positive autoregressive effects reflect the extent to which a given measure at time point $t+1$ can be predicted by the prior measure of the same construct at time point t; negative effects are usually conceptualized as a feedback system whereby the system cycles between low and high values (cf. Wright et al. 2019). Figure 1 summarizes all possible SEM paths for negative teacher feedback (NTF) and student motivation (SM).

GIMME starts with estimating an empty model across each individual, and SEM modification indices are used to indicate the anticipated improvement in model fit. To distinguish between sample, shared, and individual effects, we used previously validated cutoff values which have been shown to reliably recover true effects (Gates and Molenaar 2012; Lane et al. 2019). To identify population effects, we used a cutoff

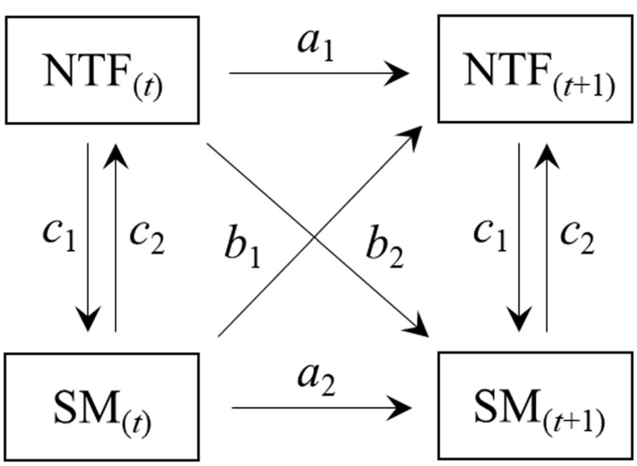

Fig. 1 First-order lagged structural model to evaluate longitudinal association patterns of negative teacher feedback (NTF) and student motivation (SM). The parameters a_1 and a_2 denote autoregressive effects, b_1 and b_2 are cross-lagged effects, and c_1 and c_2 are contemporaneous effects

value of > 75%, that is, GIMME (iteratively) identifies paths that optimally improve the fit of the group model for more than 75% of the individuals until no further paths can be found. Then, the obtained group model is pruned by eliminating those paths which, because of the freeing up of paths at later iterations, no longer are acceptable according to the 75% criterion (Gates and Molenaar 2012). Next, GIMME computes matrices based on group-level path estimates of pairs of individuals and uses a robust community detection algorithm known as Walktrap (Pons and Latapy 2006) to identify shared temporal effects (i.e., subgroup-level paths). We used a cutoff value of > 50% to identify subgroup-level paths, that is, a path had to significantly improve model fit for at least 50% of a subgroup to be identified as a shared effect. A subgroup-level pruning procedure was used, where paths were removed which are no longer significant for > 50% of individuals. Finally, during the individual-level search, significant individual paths (using $\alpha = 0.01$ for testing the significance of modification indices) were added until an "excellent" model fit was obtained, that is, when two out of the four model fit indices (root mean square error of approximation (RMSEA), non-normed fit index (NNFI), comparative fit index (CFI), standardized root mean square residual (SRMR)) suggested an excellent model fit (RMSEA < 0 .05, SRMR < 0.05, CFI > 0.95, NNFI > 0.95; Brown 2006). Because GIMME is limited to first-order VARs, lagged effects beyond previous measurement occasions (such as $\text{NTF}(t) \rightarrow \text{NTF}(t+2)$ or $\text{SM}(t) \rightarrow \text{SM}(t+2)$) were not considered. Time series data for teacher feedback and student motivation were de-trended prior model estimation (Lane et al. 2019). Full information maximum likelihood (FIML) estimation was used to accommodate missing data. Instead of imputing missing values or conducting listwise deletion, FIML uses all of the available data to estimate model parameters (Enders 2010). All analyses were performed in R (R Core Team 2019) using the package `gimme` (Lane and Gates 2017).

Results

GIMME Results

As expected, the two autoregressive paths $\text{NTF}(t) \rightarrow \text{NTF}(t+1)$ and $\text{SM}(t) \rightarrow \text{SM}(t+1)$ were identified as population effects and freely estimated for each individual. We observed a median regression weight of −0.01 ($Q_1 = -0.18$, $Q_3 = 0.15$) for $\text{NTF}(t) \rightarrow \text{NTF}(t+1)$ and −0.03 ($Q_1 = -0.16$, $Q_3 = 0.21$) for $\text{SM}(t) \rightarrow \text{SM}(t+1)$. Only 48% of the estimated autoregressive effects of NTF and 47% of autoregressive effects of SM were positive indicating relative inconsistency across days. In the present sample, negative teacher feedback and student motivation can be characterized as feedback systems cycling between low and high values.

No subgroup (shared) effects were identified using a 50% cutoff value. Instead of further lowering the subgroup detection cutoff, we focused on a substantive classification of individual-level effects that were detected for 24 students (41.4%; for the remaining 34 students only autoregressive effects, NTF(t) → NTF(t + 1) and SM(t) → SM(t + 1), were needed). Individual-level effects confirmed that student motivation and negative teacher feedback influence each other over time, but there is considerable heterogeneity in these carryover effects (cf. Fig. 2). Inspecting individual-level models reveal interesting longitudinal dynamics. Some individuals have lagged and/or contemporaneous causal links between NTF and SM, but the paths can go into opposite causal directions. For example, 7 individuals (12.1%) had a contemporaneous effect of the form SM(t) → NTF(t) (Fig. 2, lower right panel) and 8 individuals (13.8%) showed a causally reversed effect NTF(t) → SM(t) (Fig. 2, lower left panel). Further, 14 individuals (24.1%) had lagged effects SM(t) → NTF(t + 1) (Fig. 2, upper right panel) and only 7 individuals (12.1%) showed causally reversed lagged effects of the form NTF(t) → SM(t + 1) (Fig. 2, upper left panel) suggesting that, in the present sample, student motivation is more likely to affect negative teacher feedback than vice versa. While high heterogeneity is observed for all individual-level effects, negative teacher feedback tends to increase student motivation on the same day but lowers student motivation on the subsequent day.

Student Mental Health Outcomes

Next, we evaluated the predictive power of individual causal classifications (NTF → SM, SM → NTF, or independence of NTF and SM) for student post-treatment mental health. While SM → NTF constitutes the dominant causal mechanism in the present sample, we focused on the reverse causal model; students are affected by negative teacher feedback. Overall,12 students (20.7%) either had lagged or contemporaneous effects of the form NTF → SM. Figure 3 (online supplement) gives the individual-level models for the 12 students and Table 1 shows descriptive statistics for pre- and post-treatment measures of teacher-rated externalization and internalization problems, and depression. Table 2 gives the results of linear regression models for mental health outcomes adjusting for gender (1 = female, 0 = male), race (1 = black,

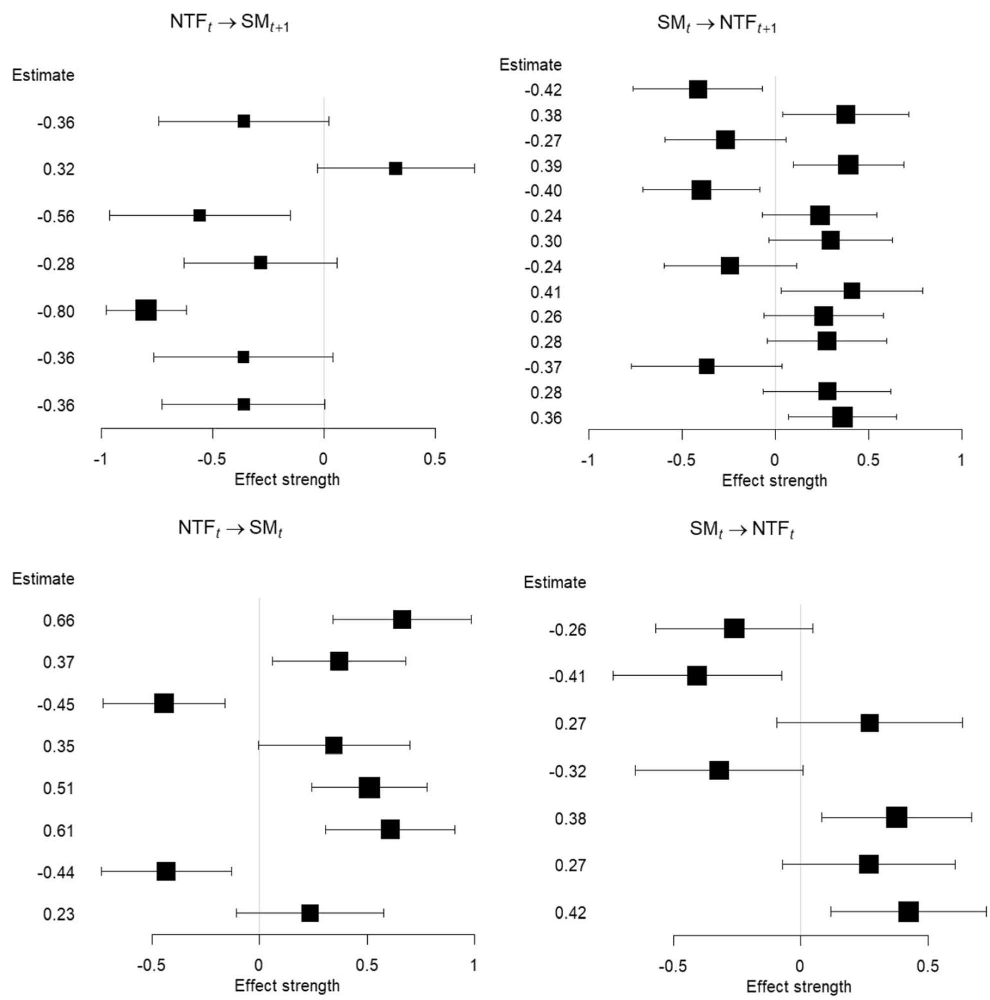

Fig. 2 Point estimates and 95% confidence intervals of GIMME individual-level effects for $n = 24$ students. The upper left panel gives the lagged effects of negative teacher feedback (NTF) on student motivation (SM; NTF(t) → SM(t + 1)), the upper right panel gives the lagged effects of student motivation on negative teacher feedback (SM(t) → NTF(t + 1)), the lower left panel summarizes contemporaneous effects of teacher feedback on student motivation (NTF(t) → SM(t)), and the lower right panel gives the contemporaneous effects of student motivation on teacher feedback (SM(t) → NTF(t))

Table 2 Multiple linear regression results for examining the effects of being influenced by negative teacher feedback (NTF → SM) on post-treatment internalization problems, externalization problems, and depression

	Internalization problems			Externalization problems			Depression		
	b	*SE*	*p*	*b*	*SE*	*p*	*b*	*SE*	*p*
NTF → SM	6.96	3.81	0.074	3.69	3.21	0.256	9.40	3.80	0.017
Baseline measure	0.46	0.12	< 0.001	0.89	0.12	< 0.001	0.71	0.12	< 0.001
Female	0.92	3.41	0.788	1.23	2.88	0.672	2.49	3.40	0.468
Black	− 1.90	3.58	0.597	− 1.57	3.00	0.603	− 4.47	3.54	0.212
FRL	3.04	3.46	0.383	2.02	3.16	0.525	2.03	3.55	0.571
SPED	6.40	3.51	0.075	0.42	2.78	0.881	2.84	3.61	0.436
Age	0.32	3.47	0.926	3.67	2.93	0.217	0.03	3.48	0.993
Cohort: 2017[(a)]	− 2.83	4.16	0.500	− 2.95	3.55	0.410	− 3.29	4.24	0.442
2018	− 0.90	3.97	0.821	− 1.77	3.41	0.606	− 1.62	4.06	0.692
R^2	0.48			0.65			0.63		

b, unstandardized regression coefficient; *SE*, standard error; *p* = *p* value; *FRL*, free/reduced lunch; *SPED*, special education service

[a] 2016 cohort is used as the reference

0 = other), age (in years), free/reduced lunch (1 = yes, 0 = no), special education (1 = yes, 0 = no), study cohort, and pretest mental health measures. Students whose motivation is affected by negative teacher feedback tend to be have more internalization problems ($b = 7.0$, $p = 0.074$, $d = 0.24$) and are more depressed ($b = 9.4$, $p = 0.017$, $d = 0.32$) at the end of the semester compared with students who are not affected by teacher feedback. There were no significant effects for externalization problems.

Finally, we examined whether baseline student-teacher relations differentially predicted the individual causal classifications (NTF → SM, SM → NTF, or independence of NTF and SM). Overall, there are no significant differences in the perceived student-teacher relation for the two GIMME groups independent of measurement occasion (pre and post intervention) and rater (student and teacher). All models were adjusted for the covariates listed above.

Discussion

We used a small subset of students participating in a study of SMARTS, a self-monitoring training program, to examine temporal dynamics between student daily ratings for their readiness or motivation for change and rates of agreement between students and teacher perceptions of goal performance feedback. Following rigorous screening and inclusion criteria, 58 students provided daily readiness for change ratings. These ratings were used alongside daily self and teacher ratings of goal performance to identify groups of individuals who appeared to fit into several patterns. Noteworthy among these patterns are subgroups of students where daily readiness ratings influenced subsequent daily performance as well as a group of students where negative teacher feedback (i.e., larger discrepancies between student self ratings and teacher ratings) adversely impacted subsequent readiness.

For most students, daily readiness begets improved alignment between student self and teacher ratings of goal performance. However, for a small subset of youth, it appears that daily readiness has an unintended or iatrogenic effect on rates of student and teacher agreement rates. One explanation of these unintended consequences is that for these students, high levels of daily readiness may contribute to them feeling overconfident and exacerbate pre-existing illusory biases. These students may benefit from daily reminders or prompts to assist them in minimizing or avoiding the overconfidence trap or overinflated perceptions of their own self and their performance abilities (Hoza et al. 2004; Owens et al. 2007).

A second subgroup effect that was observed is that if the less common causal relation $NTF(t) \rightarrow SM(t+1)$ exists for a student, there tends to be a negative relation (6 out of 7 effects were negative). In other words, for this subset of student, increased disagreement led to lowered readiness for change the next day—and in turn may have the impact of disengaging a student from an intervention where formative feedback is provided frequently. In short, and contrary to the literature on formative daily feedback, some students may require that feedback be tailored in certain ways that optimizes their capacity to internalize that feedback and contribute to positive outcomes for that student. If the purpose of ongoing performance feedback is to (a) verify whether a gap exists between current and expected performance levels, (b) learn the magnitude of that gap, (c) elaborate strategies to reduce the gap, and (d) monitor the effectiveness of those strategies (Shute 2008), then the findings here suggest that some students may require tailored feedback in ways that they can take advantage of the feedback. The manner that feedback is discussed or presented to students may be a parallel mechanism worth exploring. In this study, although we provided training and support for school personnel in presenting feedback, we were measuring variation in the presentation style of feedback discussions. Consistent with MI, the language adults use in discussing feedback with youth will influence subsequent motivation. The findings imply that more rigorous training and/or more structured feedback scripts and protocols may be helpful in minimizing the number of students who present this iatrogenic pattern.

Studies suggest teachers using task-specific and constructive feedback can positively influence student behaviors (Sutherland et al. 2000); however, findings here imply some students experience feedback differently. Our findings suggest that some students may be vulnerable to negative feedback in that it interferes with subsequent motivation and may exacerbate depressive symptoms. Depressive symptoms are commonly rooted in interpretative biases and distortions (Herman et al. 2013); thus, it is possible that students vulnerable to this effect may have cognitive predispositions to internalize negative feedback as evidence of failure or disappointment. Future research will need to include baseline measures of cognitive biases to test whether pre-existing cognitive style predicts vulnerability to iatrogenic response to feedback; if so, such students may benefit from cognitive restructuring interventions and/or feedback may need to include direct conversations to interpret negative feedback in more constructive ways.

That student-rated student-teacher relations did not predict classification groupings was surprising. This may indicate that the negative impact of teacher feedback the prior day on student readiness the next day is not due to preexisting differences in student-teacher relations but developed independently. The finding is consistent with the idea that feedback operates at the person level (i.e., internalized) so that it is not contingent on prior or subsequent relationship status. From an MI perspective, one possibility is that the more proximal influence over motivation is the way the feedback is presented by teachers to students may elicit sustained talk and that feedback leads students to be less motivated—and in turn more depressed or less happy.

There are noteworthy limitations to the study that we will attempt to address here. First and foremost, the analysis here relied on a subsample of a larger study. There is no control group in this study. The GIMME results presented here and the groupings are idiographic and relative to these students, in this study, using these measures. Further, intensive longitudinal data are known to be prone to missingness. The present results are based on 58 students (45% of the total sample), and we cannot rule out potential selection biases. The number of available time points per student was rather small (median time series length of 37.5; $Q_1 = 32.0$, $Q_3 = 38.0$) compared with previous applications of GIMME. Using Monte Carlo simulations, Lane et al. (2019) showed that, for 30 observations per subject and 5 variables, GIMME achieves a path recovery of 82.4%. Path recovery tends to increase with smaller numbers of variables, thus we can expect adequate path recovery for 2 variables (student motivation and teacher feedback). Although we were able to detect population- and individual-level effects, the present study may have been underpowered to rule out shared (sub-group) effects. We also recognize that there are many factors at play in shaping anyone's motivation—and very few of these factors are captured here. For example, some kids getting feedback about discrepancy may undermine readiness. In particular, we did not monitor the manner in which students were provided feedback. It is entirely reasonable to assume that a teacher who is exposed to MI strategies and uses the concepts of motivation to frame communications with students will have an entirely different outcome. Though this assumption can be measured and tested—and it should be—it also implies something should be done to reduce negative effects and reiterates our point; not all persons receive feedback in the same manner. If a student is involved in a daily ongoing performance feedback intervention, and there is a discrepancy between student and teacher views, then we need to be mindful about how that discrepancy is discussed. One possible factor not represented in these analyses is that we did not capture the ability of teachers and student support personnel to give feedback consistent with the values and themes of MI. Such training is vital—so is measuring the capacity to provide feedback promoting youth to engage in adaptive actions. Lastly, though we relied on similar but adapted (i.e., changed wording to read "ready" instead of "motivated") measures to reflect the conceptualization of readiness as prior studies have done (Miller and Rollnick 1991), we recognize that these measures do not fully capture the complex nature of motivation.

Future applications of GIMME hold promise for unpacking person-specific treatment responses and guiding improvements in preventive interventions that minimize iatrogenic effects. Such applications will require prevention studies to collect rich, repetitive data streams of key intervention processes most likely related to outcomes. In turn, advances in efficient, valid measures that are sensitive to change and that natural implementers can consistently use as intended are needed to best capitalize on the benefits of analytic tools such as GIMME.

Funding Information Preparation of this manuscript was supported by a grant from the US Department of Education, Institute of Education Sciences (#R305A150517).

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no other conflict of interest.

Ethical Statement All procedures performed in studies involving human participants were in accordance with ethical standards of the institutional and national research committee and the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

References

Bandura, A. (2005). The primacy of self-regulation in health promotion. *Applied Psychology, 54*, 245–254.

Beltz, A. M., & Gates, K. M. (2017). Network mapping with GIMME. *Multivariate Behavioral Research, 52*, 789–804.

Bergman, L. R., & Magnusson, D. (1997). A person-oriented approach in research on developmental psychopathology. Development and Psychopathology, 9, 291–319.

Boudreaux, E. D., Sullivan, A., Abar, B., Bernstein, S. L., Ginde, A. A., & Camargo, C. A. (2012). Motivation rulers for smoking cessation: A prospective observational examination of construct and predictive validity. *Addiction Science & Clinical Practice, 7*, 8.

Boutelle, K. N., & Kirschenbaum, D. S. (2012). Further support for consistent self-monitoring as a vital component of successful weight control. *Obesity Research, 6*, 219–224.

Bowen, N. K. (2010). Child-report data and assessment of the social environment in schools. *Research on Social Work Practice, 20*, 1–11. https://doi.org/10.1177/1049731510391675.

Brown, T. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford.

Butryn, M. L., Phelan, S., Hill, J. O., & Wing, R. R. (2012). Consistent self-monitoring of weight: A key component of successful weight loss. *Obesity, 15*, 3091–3096.

Cleary, T. J., & Zimmerman, B. J. (2004). Self-regulation empowerment program: A school-based program to enhance self-regulated and self-motivated cycles of student learning. *Psychology in the Schools, 41*, 537–550.

Crone, D. A., Hawken, L. S., & Horner, R. H. (2010). Responding to problem behavior in schools: The behavior education program. Guilford Press

Cunningham, R. M., Chermack, S. T., Zimmerman, M. A., Shope, J. T., Bingham, C. R., Blow, F. C., & Walton, M. A. (2012). Brief motivational interviewing intervention for peer violence and alcohol use in teens: One-year follow-up. *Pediatrics, 129*, 1083–1090.

Deci, E. L., & Ryan, R. M. (2011). Self-determination theory. *Handbook of theories of social psychology, 1*, 416–433.

Deno, S. L., Reschly, A. L., Lembke, E. S., Magnusson, D., Callender, S. A., Windram, H., & Stachel, N. (2009). Developing a school-wide progress-monitoring system. *Psychology in the Schools, 46*, 44–55.

Dowdy, E., Ritchey, K., & Kamphaus, R. W. (2010). School-based screening: A population-based approach to inform and monitor children's mental health needs. *School Mental Health, 2*, 166–176.

Enders, C. K. (2010). *Applied missing data analysis*. New York: Guilford Press.

Flannery, K. B., Fenning, P., Kato, M. M., & McIntosh, K. (2014). Effects of school-wide positive behavioral interventions and supports and fidelity of implementation on problem behavior in high schools. *School Psychology Quarterly, 29*, 111.

Fuchs, L. S., & Fuchs, D. (1999). Monitoring student progress toward the development of reading competence: A review of three forms of classroom-based assessment. School Psychology Review, 28.

Fuchs, L. S., Fuchs, D., Compton, D. L., Bryant, J. D., Hamlett, C. L., & Seethaler, P. M. (2007). Mathematics screening and progress monitoring at first grade: Implications for responsiveness to intervention. *Exceptional Children, 73*, 311–330.

Gates, K. M., & Molenaar, P. C. M. (2012). Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples. *NeuroImage, 63*, 310–319. https://doi.org/10.1016/j.neuroimage.2012.06.026.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica, 37*, 424–438. https://doi.org/10.2307/1912791.

Herman, K., Reinke, W., Frey, A., & Shepard, S. (2013). Motivational interviewing in schools: Strategies for engaging parents, teachers, and students. Springer Publishing Company.

Hoza, B., Gerdes, A. C., Hinshaw, S. P., Arnold, L. E., Pelham Jr., W. E., Molina, B. S., et al. (2004). Self-perceptions of competence in children with ADHD and comparison children. *Journal of Consulting and Clinical Psychology, 72*, 382.

Kaiser, B., & Rasminsky, J. S. (2009). Challenging behavior in elementary and middle school. Upper Saddle River, NJ: Pearson.

Kamphaus, R., & Reynolds, C. (2007). Behavioral & emotional screening system. NCS Pearson.

Kim, J., Zhu, W., Chang, L., Bentler, P. M., & Ernst, T. (2007). Unified structural equation modeling approach for the analysis of multisubject, multivariate functional MRI data. *Human Brain Mapping, 28*, 85–93. https://doi.org/10.1002/hbm.20259.

LaBrie, J. W., Quinlan, T., Schiffman, J. E., & Earleywine, M. E. (2005). Performance of alcohol and safer sex change rulers compared with readiness to change questionnaires. *Psychology of Addictive Behaviors, 19*, 112.

Lane, S. T., & Gates, K. M. (2017). Automated selection of robust individual-level structural equation models for time series data. *Structural Equation Modeling, 24*, 768–782. https://doi.org/10.1080/10705511.2017.1309978.

Lane, S. T., Gates, K. M., Pike, H. K., Beltz, A. M., & Wright, A. G. C. (2019). Uncovering general, shared, and unique temporal patterns in ambulatory assessment data. *Psychological Methods, 24*, 54–69.

López-Viets, V., Walker, D. D., & Miller, W. R. (2002). What is motivation to change? A scientific analysis. In M. McMurran (Ed.), *Motivating offenders to change: A guide to engagement in therapy. Chichester: WileyLütkepohl, H. (2005)*, *New introduction to multiple time series analysis*. Berlin: Springer.

Lütkepohl, H. (2005). New introduction to multiple time series analysis. Berlin: Springer.

McNamara, E. (2009). *Motivational interviewing: Theory, practice and applications with children and young people*. Ainsdale: Merseyside.

Miller, W. R., & Rollnick, S. (1991). *Preparing people to change addictive behavior*. New York: Guilford Press.

Miller, W. R., & Rollnick, S. (2012). *Motivational interviewing: Helping people change*. New York: Guilford Press.

Miller, W. R., & Rose, G. S. (2009). Toward a theory of motivational interviewing. *American Psychologist, 64*, 527.

Molenaar, P. C. (2019). Granger causality testing with intensive longitudinal data. *Prevention Science, 20*, 442–451. https://doi.org/10.1007/s11121-018-0919-0.

Owens, J. S., Goldfine, M. E., Evangelista, N. M., Hoza, B., & Kaiser, N. M. (2007). A critical review of self-perceptions and the positive illusory bias in children with ADHD. *Clinical Child and Family Psychology Review, 10*, 335–351.

Pons, P., & Latapy, M. (2006). Computing communities in large networks using random walks. *Journal of Graph Algorithms and Applications, 10*, 191–218. https://doi.org/10.7155/jgaa.00124.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for statistical computing, Vienna, Austria. URL https://www.R-project.org/.

Reinke, W. M., Lewis-Palmer, T., & Martin, E. (2007). The effect of visual performance feedback on teacher behavior-specific praise. *Behavior Modification, 31*, 247–263.

Reinke, W. M., Lewis-Palmer, T., & Merrell, K. (2008). The classroom check-up: A classwide teacher consultation model for increasing praise and decreasing disruptive behavior. *School Psychology Review, 37*, 315.

Reinke, W. M., Herman, K. C., & Sprick, R. (2011). Motivational interviewing for effective classroom management: The classroom check-up. Guilford press.

Reynolds, C. R., Kamphaus, R. W., & Vannest, K. J. (2011). Behavior assessment system for children (BASC). Encyclopedia of clinical neuropsychology, 366–371.

Shute, V. (2008). Focus on formative feedback. *Review of Educational Research, 78*, 153–189.

Sinadinovic, K., Berman, A. H., Hasson, D., & Wennberg, P. (2010). Internet-based assessment and self-monitoring of problematic alcohol and drug use. *Addictive Behaviors, 35*, 464–470.

Sutherland, K. S., Wehby, J. H., & Copeland, S. R. (2000). Effect of varying rates of behavior-specific praise on the on-task behavior of students with EBD. Journal of Emotional and Behavioral Disorders, 8, 2-8.

Thompson, A. (2014). Randomized trial of the self-management training and regulation strategy (STARS). *Research on Social Work Practice, 24*, 414–427.

Thompson, A., & Webber, K. (2010). Realigning student and teacher perceptions of school rules: A behavior management strategy for students with challenging behaviors. *Children & Schools, 32*, 71–79.

Wentzel. (2002). Are effective teachers like good parents? Teaching styles and student adjustment in early adolescence. *Child Development, 73*, 287–301.

Wiedermann, W., Bergman, L. R., & von Eye, A. (2016). Development of methods for person-oriented research. *Journal for Person-Oriented Research, 1-2*, 1–4.

Wigfield, A., & Cambria, J. (2012). Achievement motivation. The Corsini Encyclopedia of Psychology, 1–2.

Wigfield, A., Eccles, J. S., Roeser, R., & Schiefele, U. (2008). Development of achievement motivation. Child and adolescent development: An advanced course, 406–434.

Wright, A. G. C., Gates, K. M., Arizmendi, C., Lane, S. T., Woods, W. C., & Edershile, E. A. (2019). Focusing personality assessment on the person: Modeling general, shared, and person specific processes in personality and psychopathology. *Psychological Assessment, 31*, 502–515. https://doi.org/10.1037/pas0000617.