



# A Meta-Analysis of Negative Feedback on Intrinsic Motivation

Carlton J. Fong<sup>1</sup> · Erika A. Patall<sup>2</sup> · Ariana C. Vasquez<sup>3</sup> · Sandra Stautberg<sup>3</sup>

Published online: 15 August 2018

© Springer Science+Business Media, LLC, part of Springer Nature 2018

## Abstract

The most prominent view in psychological theory has been that negative feedback should generally have a detrimental impact on intrinsic motivation. Competing perspectives and caveats on this view have suggested that negative feedback may sometimes have neutral or even positive effects. This meta-analysis of 78 studies examined the effect of negative feedback on intrinsic motivation in both child and adult samples. Results indicated that negative feedback had no effect on intrinsic motivation when compared to neutral or no feedback. When compared to positive feedback, negative feedback decreased intrinsic motivation. Moderator tests revealed that the effect of negative feedback seemed to be less demotivating when (a) the feedback statement included instructional details on how to improve, (b) criterion-based standards were used to provide feedback, and (c) feedback was delivered in-person. Implications for future research and applications to real-world settings are discussed.

**Keywords** Negative feedback · Criticism · Praise · Intrinsic motivation · Perceived competence · Meta-analysis

In 1911, Thorndike (1911) published *Animal Intelligence*, examining the influence of feedback on behavior. Presently, the notion of feedback is as critical as it was a century ago, if not more

---

Parts of this paper were presented at the annual meeting of the American Educational Research Association and biennial meeting of the Society for Research on Adolescence. This project has been funded through the American Psychological Association, the American Educational Research Association, and the Association for Psychological Science. Part of this work was from Carlton Fong's doctoral dissertation, and the authors would like to thank the committee members: Diane L. Schallert, David S. Yeager, Jennifer H. Corpus, and Marilla D. Svinicki.

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s10648-018-9446-6>) contains supplementary material, which is available to authorized users.

---

✉ Carlton J. Fong  
carltonfong@txstate.edu

<sup>1</sup> Department of Curriculum and Instruction, Texas State University, San Marcos, TX 78666, USA

<sup>2</sup> Rossier School of Education, University of Southern California, Los Angeles, USA

<sup>3</sup> Department of Educational Psychology, University of Texas at Austin, Austin, USA

so in light of high-stakes assessments and mandated learning gains in an age of accountability (Harris et al. 2014). Feedback was and still is one of the most powerful influences on learning (Hattie 2012; Van der Kleij et al. 2015), yet the question remains: What are effective ways of providing feedback to enhance intrinsic motivation?

On the whole, negative feedback is believed to be an unavoidable practice when providing a performance evaluation, in spite of its often ego-threatening consequences (Dahling and Ruppel 2016). Ilgen and Davis (2000) argued “few beliefs are more widely accepted by psychologists, managers, educators, and others concerned with human performance than the belief that people need to receive feedback about how well they are performing their tasks/jobs” (pp. 550–551). However, theorists and educators view negative feedback as a conundrum or a dilemma (Van-Dijk and Kluger 2004). In the classroom, giving information that highlights shortcomings in a student’s work can simultaneously instruct the student toward greater gains in learning, yet undermine self-confidence and intrinsic motivation (IM) or the proclivity to engage in an activity due to the inherent satisfaction it brings (Cohen et al. 1999). Teachers, mentors, parents, employers, and coaches often struggle to provide negative feedback in a motivating way (Yeager et al. 2014).

The focus of our review was on intrinsic motivation, first because, intrinsic motivation has been found to predict higher learning, well-being, and psychosocial functioning (Cerasoli et al. 2014; Mabbe et al. 2018; Ryan and Deci 2000). Given the various educational advantages associated with intrinsic motivation, understanding the antecedents of IM, such as various forms of feedback, is critical. Second, although knowledge of how negative feedback affects task performance is also important, a prior review on the topic (Kluger and Denisi 1996) yielded confusing results. They found that feedback interventions had a positive effect on performance overall, but the moderation of feedback valence was not significant. The authors said this lack of significant moderation was conspicuous and could not identify any theoretical explanation for how feedback valence did not affect performance. Instead, they suggested that meta-task processes, such as motivation and personality, be examined. In light of these issues, we reviewed prior literature in the fields of education and psychology and conducted a meta-analysis to determine the overall effect of negative feedback on IM and perceived competence and the factors that explain variation in this relationship.

## The Role of Feedback Valence in Intrinsic Motivation

Feedback is generally understood as the numerous procedures that are used to tell a learner if a response is right or wrong (Kulhavy 1977). It is inherently a response to one’s performance or understanding (Hattie and Timperley 2007). Duijnhouwer et al. (2012) described feedback as “information provided by an external agent regarding some aspect(s) of the learner’s task performance, intended to modify the learner’s cognition, motivation, and/or behavior for the purpose of improving performance” (p. 171).

**Defining Negative Feedback** One of the many factors that needs to be considered when giving and receiving feedback is valence or sign (i.e., positive or negative; Kulhavy 1977). Some feedback may communicate positive qualities about a product, action, or person. Alternatively, feedback may convey neutral or negative qualities. For the purposes of this research synthesis, we defined *negative feedback* as negative evaluations made by a person of another’s products, performances, or attributes, where the evaluator presumes the validity of

the standards on which the evaluation is based. This definition provides a counterpart to the definition of positive feedback offered by praise researchers (Henderlong and Lepper 2002; Kanouse et al. 1981). However, not all negative feedback is created equal. Although negative feedback may inherently contain a message of unsatisfactory performance that threatens intrinsic motivation, a wide range of other elements within a negative feedback message may ultimately benefit students' motivation. For the scope of this synthesis, we focus on feedback that provides a negative evaluation (e.g., criticism) as the intervention of interest rather than all forms of feedback more broadly. That being said, the types of feedback that negative feedback are being compared to in our meta-analysis may range from positive feedback to no evaluation whatsoever. In learning contexts, students more often than not will receive feedback that evaluates their performance. This process is often followed by a complex set of emotions and motivations (Fong et al. 2016). Thus, we aimed to capture these motivational shifts in our study. Moreover, we wanted to make it clear that the experiments identified in the meta-analysis manipulated negative feedback as non-contingent on participant performance.

**Intrinsic Motivation and Feedback** Intrinsic motivation (IM) is the propensity to engage in a task out of interest or enjoyment, for its own sake, or without any external incentive or reward (Deci and Ryan 1985a). Feedback is believed to be a significant factor in fostering intrinsic motivation (Lepper and Chabay 1985). Analyzing the effects of positive feedback on IM, two research syntheses (Deci et al. 1999; Henderlong and Lepper 2002) showed that overall, positive feedback enhances IM, as it affirms a sense of competence. However, a synthesis on the influence of *negative feedback* on intrinsic motivation has yet to be conducted, despite 40 years of research having accumulated on the topic. In fact, mixed evidence regarding the effect of negative feedback on IM has led to uncertainty regarding both the direction and magnitude of its effect, making a meta-analysis of the topic particularly timely (Bracken et al. 2004; Elliot et al. 2000). Van-Dijk and Kluger (2004) corroborated this issue, “despite our common sense notion that indicates that feedback sign (positive vs. negative) has a decisive effect on motivation, the vast literature has no clear specifications when and how positive (negative) feedback increases or decreases motivation” (p. 113).

Understanding the differences between positive and negative feedback on students' IM would indeed be fruitful (Voerman et al. 2015). Because of the manifold benefits of intrinsic motivation on academic achievement and psychological functioning, IM is an important outcome in its own right (Cerasoli et al. 2014; Mabbe et al. 2018). Moreover, the effect of negative feedback on IM is vital to the extent that IM can also influence performance; this is one possible pathway among many through which negative feedback may influence performance. As such, it deserves to be better understood as part of a larger effort to disentangle the complex effects of negative feedback on performance. Therefore, identifying the degree to which negative feedback may enhance or thwart IM and the particular conditions, these effects may be altered will benefit both theory and practice. Although feedback may also have an impact on extrinsic motivation, or the propensity to engage in a behavior because it leads to a separable outcome (e.g., praise or money; Ryan and Deci 2000), the body of literature focused on extrinsic motivation is limited. As such, this meta-analysis will address negative feedback effects on IM. As a secondary outcome of interest, we will also address perceived competence as this construct is closely linked with how feedback influences IM (discussed in the following sections).

The most prominent motivation theories suggest that negative feedback has an overall negative effect on intrinsic motivation. Self-determination theory (Deci and Ryan 2016) suggests that on average, negative feedback decreases motivation. On the other hand, two other learning and motivation theories suggest that negative feedback has the potential to increase intrinsic motivation: information processing theory (Mayer 1996) and goal-setting theory (Locke and Latham 1990). The following section highlights two contrasting views that negative feedback may either enhance or undermine IM. Although we begin by broadly discussing negatively valenced feedback helping or hindering IM, this is followed by a more nuanced distinction of negative feedback characteristics that shape students' perceptions of the evaluation.

## Negative Feedback Undermines Intrinsic Motivation

According to some scholars, in spite of the best intentions to improve subsequent performance, negative feedback may most typically produce the opposite intended effect (Ilgen and Davis 2000). Fundamentally, negative feedback's deleterious effects can be explained by its necessarily evaluative nature toward the self (Hu et al. 2016). Henderlong and Lepper (2002) argued that being evaluated engenders a contingent sense of worth that can lead to increased self-consciousness and undermine IM. A student's tendency toward self-enhancement, or the desire to elevate one's self-concept and protect the self from negative evaluation, is nearly axiomatic (Elliot et al. 2000). This suggests that negative feedback would threaten students' IM.

Self-determination theory (Ryan and Deci 2000) provides one of the most comprehensive frameworks for understanding how feedback influences IM. In particular, a sub-theory of self-determination theory (SDT), cognitive evaluation theory, posits that three fundamental needs underlie intrinsic motivation: competence, autonomy, and relatedness (Ryan and Deci 2000). Competence refers to perceived effectiveness in dealing with the environment in which a person is situated. Autonomy is the sense that one is the origin of his or her own actions. Relatedness, also referred to as belongingness, is the experience of being connected with and engaging in mutual care with others (Niemic and Ryan 2009). SDT posits that social contexts that satisfy these needs will enhance IM. Likewise, conditions that thwart satisfaction of these needs will diminish IM. As such, feedback may be understood as a vital motivational input and one of the primary ways of supporting or diminishing IM through psychological need satisfaction, namely autonomy and competence (Burgers et al. 2015). On the one hand, knowledge that one is competent and the desire to be competent are understood as pathways through which positive feedback can enhance intrinsic motivation (Elliot et al. 2000). However, negative feedback may damage one's perceptions of success and signal incompetence, consequently diminishing one's IM (Deci and Cascio 1972).

Despite self-determination theory's theoretical prediction that negative feedback will generally undermine IM because of its detrimental effect on competence beliefs, not all negative feedback is expected to diminish competence beliefs. SDT researchers describe competence-supportive or effectance-relevant feedback as including some praise and information on how one can improve in the task or in their goal pursuit (Deci and Ryan 2016). Effectance-relevant feedback, or what we call *instructional feedback* in our study, provides behaviorally relevant information and has been shown to increase IM (Deci et al. 1999). However, this informational aspect has only been primarily studied with praise.

Regarding autonomy need satisfaction, negative feedback may diminish IM to the extent that people often perceive feedback and how to respond to it as something out of their control, subsequently thwarting experiences of autonomy (De Muynck et al. 2017). Moreover, negative feedback may inherently call attention to the controlling behavior of evaluators, thereby shifting an individual's perceived locus of causality from being more internal to more external and, in turn, dampening IM (Henderlong and Lepper 2002). Said a different way, feedback may reduce an individual's sense of autonomy and subsequent IM because it is often perceived as controlling and intended primarily to coerce behavior meeting an externally imposed standard.

## Negative Feedback Enhances Intrinsic Motivation

In general, goal-setting theorists and constructivist theorists support that negative feedback may impact one's need for competence, which in turn can enhance IM. According to Butler and Winne (1995), feedback inherently catalyzes behavior as students self-regulate their learning. Goal theorists posit that people require feedback that reveals progress in relation to their goals (Fishbach and Finkelstein 2012). After people set goals, negative feedback signals discrepancies between what they have achieved and what they wish to achieve. In response, self-dissatisfaction is created and serves as a motivational inducement toward greater effort. Without such formative evaluation, students cannot adjust the level or direction of their effort to achieve their goals. Self-rated exertion of mental effort was higher after receiving negative feedback compared to positive or no feedback (Raaijmakers et al. 2017). In sum, negative feedback may motivate students to keep congruence between goals and behaviors (Cianci et al. 2010), so their needs for competence are met. Ryan et al. (1983) summarized this concept well: "A person needs to be able to get some sense of how well he or she is doing at the activity to remain intrinsically motivated" (p. 738).

From the information processing or constructivist perspective, feedback is understood as information for a learner, in the form of errors that need to be corrected and increased awareness of one's level of understanding (Ilies et al. 2010). Upon receiving negative feedback, a learner must interpret this source of instruction and either accept, modify, or reject the feedback (Kulhavy and Stock 1989). Negative feedback provides an opportunity to correct mistakes and improve task performance, filling a gap between what is understood and what is aimed to be understood (Sadler 1989; Williams et al. 2010). Butler and Winne (1995) summarized "feedback is information with which a learner can confirm, add to, overwrite, tune, or restructure information in memory, whether that information is domain knowledge, meta-cognitive knowledge, beliefs about self and tasks, or cognitive tactics and strategies" (p. 275; see Alexander et al. 1991).

Together, these theories espouse how feedback can influence one's goals, effort, and information processing, and rather as a means than an end, these in turn can influence an individual's need for competence. To the extent that feedback motivates greater effort to reach a goal or increased levels of information processing to learn, one's need for competence can be met and thereby increase feelings of self-determination (Hirst 1988). That is, negative feedback may enhance IM through the satisfaction of competence needs, which is facilitated by increased goal pursuit, exertion of effort, and information processing.

In sum, there is both empirical evidence and theoretical discussion to support both views of the effect of negative feedback. In some cases, negative feedback may improve IM, as

feedback serves as useful information to enhance one's competence and thus feelings of being self-determined. In other cases, negative feedback may be detrimental to IM, as it damages one's self-view and sense of perceived competence. Moreover, another theoretical inconsistency exists in the literature regarding the desire for positivity versus subjective accuracy (Taylor and Brown 1988). Some theorists have argued that accurate perceptions of the self are essential for one's well-being (Ayduk et al. 2013; Swann Jr 2011). At the same time, research evidence supports the pancultural desire for positive self-evaluation and perceptions of mastery and control (Sedikides et al. 2003). Thus, whether negative feedback enhances one's self-accuracy or damages needs for self-enhancement as it relates to intrinsic motivation needs to be resolved.

In the following section, we will attempt to address additional factors that are likely to determine the conditions under which the effects of negative feedback on motivation may vary.

## Factors that May Influence the Effects of Negative Feedback on Intrinsic Motivation

It seems reasonable to expect that the effect of negative feedback on IM may change depending on various aspects and circumstances. Characteristics of the feedback, task, evaluator, feedback receiver, and research design may all influence the effect of negative feedback. In the following sections, we describe each of the theoretically relevant moderators to the feedback-motivation relationship.

### Characteristics of the Feedback

Theory and empirical evidence underscore the importance of both the content and delivery of a negative feedback statement. In the subsequent sections, we will describe the following characteristics of feedback content: normative comparisons versus objective criteria; attributing performance outcomes to one's level of effort ability; and the inclusion of praise or information about how to improve. Likewise, we will also highlight various aspects of the manner negative feedback can be delivered: in a public versus private context; using autonomy-supportive versus controlling language; and degree of threat. Examination of these features challenges the simplistic view that negative feedback is harmful to students' IM in its messaging because it carries an unfavorable evaluation. Because negative feedback can vary along all the dimensions previously discussed, the term *negative feedback* is perhaps too crude to understand the many nuances the literature has described. With this in mind, we discuss how a number of aspects can sharpen our understanding of the *negative feedback* construct.

### Normative Versus Criterion-Based Feedback

Research has sometimes found positive feedback that focuses on social comparison or normative standards (e.g., "Good job, you scored higher than 80% of your peers.") to be more enjoyable compared to receiving no praise at all (Gaines et al. 2005). In contrast, other studies suggest that objective or criterion-based praise (e.g., "You answered 10 out of 10 correctly.") will be more beneficial than social-comparison praise (Butler 1987). In particular, an overreliance on normative positive feedback can lead to decreased persistence during setbacks (Corpus et al. 2006). Corpus et al. (2006) argued that normative praise may prevent

children from enjoying a task. Relying on social comparisons when providing feedback inadequately equips students during situations when they are outperformed.

The comparison of normative and criterion-based *negative* feedback has rarely been examined in primary experiments. The current meta-analysis provides an opportunity to compare the effects of normative negative feedback (e.g., “You did worse than the majority of your classmates.”) with that of criterion-based negative feedback (e.g., “8 out of your 10 answers were incorrect.”) across all included studies. In the face of challenge, criterion feedback focuses the feedback receiver on the task, rather than the performance of others. As such, improvement may appear more controllable with criterion feedback. Individuals may have some control over their own performance, but little control over other people. Thus, we predicted that normative negative feedback would be more deleterious for IM than criterion-based negative feedback.

### **Attributional Feedback**

Whether feedback includes information that attributes performance to ability versus effort is likely to moderate its relation with IM (Fishbach and Finkelstein 2012). The reason being that the two attribution targets of effort and ability are generally perceived to vary in their stability and controllability, and thus, implications for improvement in the future (Weiner 2014). Feedback that attributes performance to ability implies that there is little an individual can do because ability remains stable. Even after the provision of positive feedback, attributing performance to ability as opposed to effort may be motivationally detrimental once individuals face a challenge or subsequent failure (Henderlong and Lepper 2002). Likewise, negative feedback is likely to be more detrimental to IM when focused on ability rather than effort. Poor performances attributed to one’s ability may perpetuate the notion that improvement is impossible. Because effort is within the individual’s control, effort-focused criticism may increase IM (Skipper and Douglas 2015).

### **Instructional Feedback**

The extent to which feedback includes instructional directions for how to improve may also moderate negative feedback’s effect on IM. Instructional feedback, or feedback geared toward task improvement, is known in the field as a variety of terms: formative feedback (Shute 2008), corrective feedback (Hattie and Timperley 2007), constructive feedback (Duijnhouwer et al. 2012; Fong et al. 2016), and effectance-relevant feedback (Deci and Ryan 2016). Instructional feedback has been defined as non-confrontational feedback that provides specific directions for improvement and is delivered with sensitivity about attributing blame (Baron 1988). It is a powerful tool for enhancing student motivation as it provides information regarding what to do and how to respond in the future (Bangert-Drowns et al. 1991; Fong et al. 2016; Phye and Sanders 1994; Shute 2008). If instructional feedback facilitates the goal of improved performance, then the need for competence is met, which thereby increases intrinsic motivation. With this in mind, we predicted instructional negative feedback would be more intrinsically motivating compared to feedback without constructive elements.

### **Inclusion of Praise or Rewards**

One suggested solution for mitigating the ego-threatening aspect of negative feedback is to include elements of praise (Yeager et al. 2014). Including praise along with criticism can boost

self-esteem and lessen the demotivating effect of negative feedback (Brummelman et al. 2014). This effect is particularly salient when feedback focuses on the process or effort that students use while engaging in a task (Kamins and Dweck 1999).

### Social Nature of Feedback

Another potentially important dimension of feedback is the social circumstances under which it is delivered: in-person versus mediated through other modes (i.e., paper, computer, audio cues). Research suggests that private, or a mediated delivery of feedback (where only the participant knows the nature of the evaluation) will have more desirable effects on IM than public or in-person delivery (at least one other person, such as an evaluator, knows about the nature of the feedback; Ames 1992). According to SDT, public delivery is expected to decrease IM because there may be greater external pressure. This, in turn, lowers one's internal locus of causality to respond to the feedback because others are aware of one's performance (Deci and Ryan 1985b). Some studies have shown that in-person criticism in which both the recipient and evaluator are aware of the negative feedback is demotivating (Bracken et al. 2004). In contrast, other research has found no difference among negative feedback delivery conditions (Lin et al. 2013).

### Autonomy-Supportive Versus Controlling Feedback

According to cognitive evaluation theory, the extent to which feedback is delivered in an autonomy-supportive versus controlling manner will influence its effect on IM (Deci et al. 1999). Negative feedback delivered with controlling language informs the individual of what he or she *should* have done or *needs* to do in the future. This controlling feedback may undermine IM more than feedback that uses non-controlling language to the extent that the controlling language is itself autonomy-thwarting (Reeve 2016). In contrast, autonomy-supportive language communicates that the individual is in control of his or her own behavior (i.e., "you could do" or "you might consider") and bolsters IM. A related form of autonomy-supportive feedback is "wise" feedback (Cohen et al. 1999; Yeager et al. 2014), which instills trust between feedback receiver and evaluator as well as sets high expectations for the student. This form of feedback has been found to be highly motivating (Yeager et al. 2014). Thus, delivering negative feedback in an autonomy-supportive manner may mitigate potential detrimental effects on IM.

### Degree of Threat in Feedback

The degree of threat within feedback statements may also alter their effects on IM. For instance, the undermining effect of negative feedback on IM may be negligible when delivered in a milder tone. Anderson and Rodin (1989) examined effects of mild negative feedback by providing normative feedback that suggested a score ranked slightly above 50th percentile. This negative feedback was considered mild to the extent that it indicated that the individual was still about average in their performance. They found that participants felt discouraged but did not perceive the feedback to diminish self-confidence compared to receiving harsher forms of feedback. Their study resulted in two important findings. First, milder negative feedback undermined IM less than harsh negative feedback. Second, it actually had a positive effect on IM compared to receiving no feedback at all.



## Characteristics of the Task

There is a lack of experimental research focused on how task characteristics influence the effect of negative feedback on IM. However, task characteristics such as the level of interestingness, difficulty, goal structure, and perceived importance may be important moderators. Although these task characteristics are theoretically relevant, empirical studies tend to only report task interestingness as it relates to IM, which will be the focus for this study. In one of the earliest SDT studies examining the effect of negative feedback, Deci and Cascio (1972) found that individuals who had received negative feedback during an interesting task had lower IM compared to individuals who received praise. In fact, it would seem reasonable that the effect of feedback may be minimal when interest in the task or task difficulty is low to begin with. That is, while engaged in an uninteresting or easy task, there are perhaps lower levels of IM to undermine in the first place. If the task is of little interest prior to receiving feedback, then any possible reduction in interest due to the negative feedback is limited and likely to be negligible (Hirst 1988).

## Characteristics of the Feedback Receiver

Next, we turn our attention to aspects of the feedback receiver that may moderate how negative feedback influences IM. Prior research has suggested that age and gender may be two important characteristics of the feedback receiver that moderate the effect of feedback on IM. Specifically, a meta-analysis conducted by Deci et al. (1999) suggested that the effect of praise was not ubiquitous across age. Verbal reinforcements enhanced IM among college students, but not among younger children. Regarding gender differences, they found that women experienced praise as more controlling than men, which led to decrements in IM after receiving positive feedback. Likewise, some research has found that women are more sensitive to negative feedback and experience greater declines in IM as a function of negative feedback (Vallerand and Reid 1988). However, these findings are not consistent across all studies, as other research has found no difference in the effects of negative feedback across genders (Shanab et al. 1981).

## Characteristics of the Evaluator

Regarding characteristics of the evaluator and their moderating influence on negative feedback and IM, feedback is likely most persuasive when the people who provide the information are viewed as knowledgeable and reliable (Bong and Skaalvik 2003). Thus, the effects of negative feedback are likely to be mitigated when the evaluator is perceived to lack credibility. Furthermore, the quality of the relationship between evaluator and feedback receiver may moderate the effect of negative feedback (Fong et al. 2018a). In the context of a close and caring relationship, feedback may be perceived as more authentic and intended to help (Henderlong and Lepper 2002). In contrast, feedback may be received as controlling if there is mistrust or poor relationship quality (Bryk and Schneider 2002), potentially leading to the dismissal of the feedback. A closer relationship may also lead the feedback receiver to perceive the evaluator as sincere; sincerity is described as a necessary condition in order for feedback to be accepted and to have a positive motivational effect (Henderlong and Lepper 2002).

## Characteristics of Research Methods

Lastly, characteristics of the research methods may influence the extent to which the effect of negative feedback on IM can be detected or even the nature of the effect. One important factor is the nature of the comparison condition, specifically whether the comparison to negative feedback is positive feedback or neutral/no feedback (Hattie and Timperley 2007). As described earlier, SDT assumes that negative feedback generally undermines IM and positive feedback enhances IM. Thus, the undermining effect of negative feedback is likely to be stronger when compared to praise rather than no feedback or neutral feedback (e.g., “You completed the task.”).

Second, another methodological issue is the type of IM measure used: self-reported or behavioral. Deci et al. (1999) found differential effects of praise depending on what kind of IM measure was included. Further, non-significant correlations between behavioral and self-reported IM measures have caused doubt on whether they index the same construct (Wicker et al. 1990). Although self-reported measures are subject to biases such as social desirability, acquiescence, and retrospective reconstruction of events, they may be more sensitive to manipulations compared to behavioral measures (Patall et al. 2008).

The most common behavioral measure of intrinsic motivation is task persistence. The extent to which the contingencies for continuing on the task when motivation is measured by a period of free choice play is largely considered an intrinsic process. That being said, when feedback motivates persistence out of a desire to improve on the task, the type of motivation experienced by the participant may not be purely intrinsic. There may be some degree of ego-involvement when responding to feedback (Ryan et al. 1991). Thus, we cannot be fully confident whether students are persisting in a task for intrinsic or other internal reasons. In spite of this limitation, we default to how task persistence is framed in the majority of the literature as intrinsically motivated.

## Present Study

The present study is guided by the following primary questions: What is the overall effect of negative feedback on intrinsic motivation as a primary outcome and perceived competence as a secondary outcome, when compared with positive feedback and neutral or no feedback? What factors explain variation in these relationships? When comparing types of negative feedback, which feedback features are most motivating? Previous research has suggested that many theoretically relevant factors may influence the effect of negative feedback on IM (e.g., Hattie and Timperley 2007), including but not limited to the social context of delivery, feedback standard, motivation features of the feedback, task interestingness, and the feedback recipient’s age and gender. We hypothesize that these factors may moderate the effect of negative feedback on IM and perceived competence.

## Method

### Inclusion Criteria

To be included in the meta-analysis, a study was required to meet several criteria. All studies needed to employ a negative feedback manipulation using random assignment on a recent task performance. This means that participants in one condition received some type of negative

feedback, and participants in the comparison group either received no, neutral, or positive feedback. The comparison condition could also be another type of negative feedback to assess whether particular elements of negative feedback may moderate its effect on IM. In addition, there were no explicit criteria set for participants, subject matter, or task, as we were interested in the effects of feedback in all learning contexts. Reports written in non-English language were excluded. Both published and unpublished literature sources were included as recommended by Polanin et al. (2016).

Because the effect of feedback on IM was of primary interest to this meta-analysis, a study had to include a measure of IM. We included any measure of task interest or enjoyment, time spent on a task without external pressure or constraints, or reports of willingness to engage in the task again in the future or task persistence. In the literature, there have been multiple ways to assess IM. The most common measures, which have been prevalent for decades, are self-reported assessments of enjoyment or behavioral assessments of task persistence. Both of these measures are not without challenges. Self-reported measures may be subject to reporting biases and do not capture the action potential on IM sufficiently. However, they provide a glimpse at an individual's subjective experience. Even in the absence of contingencies following a task, willingness to engage and task persistence could be a reflection of non-intrinsic in addition to intrinsic sources of motivation, but they provide a behavioral indicator of IM. Although some have challenged if both these measures capture intrinsic motivation (Ryan et al. 1991), we have included both given the history and prevalence of these measures in prior studies and later assess any differences in a moderator test. As a secondary outcome of interest, within our pool of included studies, we also coded perceived competence outcomes and included them in our meta-analysis. As discussed in our theoretical framework, perceived competence plays an important role in how feedback may influence intrinsic motivation. Note that we only included perceived competence if the study measured both intrinsic motivation and perceived competence. Thus, this is not necessarily an exhaustive review for the perceived competence outcome.

If a study did not report an IM measure, it was excluded. Despite interest in how feedback may influence task performance, this outcome was not included in this meta-analysis due to a prior review by Kluger and Denisi (1996). Although this review is over two decades old, and more studies on the topic would have been conducted, we argue that such an update on this outcome would not necessarily yield different results. Since this review, one trend in the literature that may provide an updated understanding of feedback effects is the emergence of feedback in computer-based environments. However, the effects of feedback on students' learning and performance in computer-based settings have also been reviewed in a recent meta-analysis by Van der Kleij et al. (2015). Thus, we argue that the impact of feedback on task performance has been sufficiently reviewed, and the more pressing need is to examine intrinsic motivation as the outcome.

## Literature Search Procedures

Multiple strategies were used to locate all possible relevant studies that met the inclusion criteria. First, the following electronic reference databases were searched for documents cataloged before January 2017: PsycINFO, ERIC (Educational Resources Information Clearinghouse), Dissertations and Theses Global, and Google Scholar. For each database, we used a string of search terms with a required keyword regarding feedback (*feedback* search terms: feedback, evaluation, critici\*, teacher response, performance information) and a required keyword regarding intrinsic motivation (*motivation* search terms: motivation, intrinsic value, interest, persistence, self-determination, autonom\*). We applied appropriate Boolean

techniques and truncation as indicated by \*, allowing for multiple forms of the keyword such as *criticism* and *criticizing* for *critici\**, and *autonomous regulation* and *autonomous motivation* for *autonom\**. Note that the keyword *motivation* subsumes terms like *intrinsic motivation*. Also, keywords could be located anywhere in the full-text to maximize the opportunities to identify relevant reports. This search strategy was designed to capture comprehensively all studies with outcomes related to IM.

Second, once this search strategy was employed, and all citations had been retrieved, abstracts for these studies were judged for relevance, resulting in a pool of studies that would possibly meet the inclusion criteria. The full texts of potentially codeable studies were reviewed and evaluated with the inclusion criteria. Third, ancestry searches were conducted by reviewing the reference section of relevant studies retained for coding. Fourth, in January 2017, descendent searches were conducted in Social Sciences Citation Index for the following two articles (Deci and Cascio 1972; Deci et al. 1973), to find papers that had cited these early pieces on the effect of negative feedback on motivation. Fifth, in January 2017, we conducted hand searches of journals that contained multiple included studies: *Journal of Personality and Social Psychology*, *Journal of Social Psychology*, *Journal of Experimental Social Psychology*, *Motivation and Emotion*, and *Journal of Educational Psychology*.

Sixth, unpublished data or gray literature were obtained through contacting listservs, once during Fall 2013 and Spring 2017, to the following organizations: Motivation in Education Special Interest Group ( $n = 425$ ) and Division C: Learning and Instruction ( $n =$  approximately 4500) from the American Education Research Association and the divisions of Educational Psychology ( $n = 1108$ ), Social Psychology ( $n = 1791$ ), and Sport, Exercise, and Performance Psychology ( $n = 718$ ) from the American Psychological Association. Based on membership totals from these organizations, an estimated number of members reached is approximately 8500 (Note that many individuals may have memberships in multiple organizations). Seventh, in Spring 2017, requests via electronic mail were made to several prominent researchers in the motivation and feedback areas regarding access to any relevant data that were not publicly available. Specifically, three researchers whom the database of studies revealed have published two or more studies on negative feedback and motivation were contacted directly in order to access research that would not be included in reference lists or citation databases.

### Information Retrieved from Studies

Numerous characteristics of each study were coded directly from the research report. In some instances, some inference was necessary such as using pre-established definitions to code ambiguous characteristics. The coded characteristics encompassed the following broad distinctions among studies: research report, research design, feedback manipulation, task, sample, evaluator, outcome measure (e.g., free time spent on task, self-reported interest level), and estimate of the effect size. Table 1 provides further description of the information retrieved from the included studies.

First, we identified peer-reviewed journal articles as published and everything else as unpublished (gray literature). Second, we were interested in examining the role of negative feedback in a variety of learning settings. However, after coding for study context, we observed that some studies with PK-12 students occurred in schools, the feedback manipulation was conducted by experimenters and not their actual instructors. There were no field experiments.

**Table 1** Description of information retrieved from studies

Code categories	Description
Research report	Journal article, dissertation/thesis, conference paper
Research design	
Setting	Laboratory, applied setting (school, athletic event, work setting)
Comparison condition	No/neutral feedback, positive feedback, negative feedback
Feedback	
Social context	Delivered in-person, mediated (written, visual, computer-based)
Normative	Compared to a social standard (“worse than 80% of class”)
Criterion	Compared to an objective score (“only 1 out of 10 correct”)
Instructional	Provided directions for improvement
Autonomy-supportive	Offered choices, expressed relevance, and acknowledged affect
Controlling	Language uses “must/should” statements, threats, and directives
Attributional	Targeted participants’ ability or effort
Inclusion of praise	Included positive content regarding performance
Level	Task, process, self-regulation, self
Quantity/Timing	How much feedback, time between task performance and feedback
Degree of threat	Included excessive unfriendly language
Wise	Openly and honestly established high standards and potential
Task	
Interestingness	Task was reported as interesting or not
Sample	
Gender	Percentage of female students reported
Age	P-12, college-aged
Ethnicity	Ethnic composition and percentage of each ethnicity reported
Country of origin	USA, non-USA
Motivation	Goal orientation, regulatory focus
Ability	Performance level
Self-Beliefs	Perceived competence
Evaluator	
Quality of relationship	Relationship between evaluator and receiver (close, distant)
Expertise	Level of expertise of evaluator
Outcome measure	
Self-reported	A single-item or scale of motivation, interest, liking, re-engagement
Behavioral	A measure of task persistence, time on task, or number of attempts
Effect size	Standardized mean difference (Hedges’ <i>g</i> )

We also coded the comparison condition to negative feedback. The comparators were positive feedback, no feedback, neutral feedback, or another form of negative feedback. These distinctions led us to separate our analyses into several sets: (a) we meta-analyzed effect sizes that compared positive with negative feedback; (b) we tested if effects comparing negative feedback to neutral feedback versus no feedback comparators were distinct from one another, and if not then, we combined them; and (c) we grouped effect sizes that compared negative feedback with another type of negative feedback.

Third, we coded if the feedback manipulation used in the study had the following characteristics: instructional aspects, autonomy-supportive or controlling language, attributions about effort and ability, degree of threat, wise versus unbuffered, task- versus process-focused, and the inclusion of praise. Controlling language consisted of using words like *should* or *must* and was known to reduce one’s sense of autonomy. *Wise* feedback (Yeager et al. 2014) involved high standards and assurance to the feedback receiver to try harder, conveying respect as an individual, and a lack of judgment in light of a negative stereotype. The *unbuffered* condition just provided feedback with some direction to improve but lacks the

establishment of high standards and assurance. Task-focused feedback that evaluates the product or performance was compared with process-focused feedback that evaluates the strategy or approach used for the task. Also, if the negative feedback message included a positive evaluation in combination with a negative evaluation, we would code the negative feedback as including praise.

We then coded for the social context of the feedback (if it was in-person or mediated by a written, visual, or audio message) and the comparison to either a normative or criterion-based standard. We were also interested in other characteristics of feedback such as the quantity, timing, and the level of feedback (task, process, self-regulation, self; Hattie and Timperley 2007). However, nearly all studies included feedback targeted at the task level. In regard to quantity and timing, the vast majority of studies had a single dose of feedback occurring immediately after task completion.

Fourth, we identified whether a task was reported as interesting or not. The fifth category of codes entailed details of the sample (gender, age, ethnicity, country of origin). These characteristics, when available, were extracted from the method section of included studies by recording the composition of the sample and the percentage of each subcategory when available (see Table 1 for more details). The sixth category consisted of characteristics of the evaluator (quality of relationship between evaluator and receiver and level of expertise); however, no studies included these attributes due to the prevalence of laboratory studies and experimenter-provided feedback. Seventh, we coded whether the outcome was a self-reported measure or an observed behavior. Eighth, we coded effect size information, which is outlined in more detail in subsequent sections.

If these categories were not present, we coded them as “not reported.” With such little variance among particular moderators, meaningful tests were unable to be conducted. Although we coded for an exhaustive list of study and sample characteristics, we want to once again reiterate that due to a lack of reporting from primary authors or variance among the included studies of that particular moderating variable, meaningful analyses could not be conducted.

## Effect Size Calculation

We used the standardized mean difference to estimate the effect of negative feedback on motivation and related outcomes. This is a scale-free measure of the distance between two group means, which is calculated by dividing the difference between two group means by a pooled standard deviation. We subtracted the comparison condition from the negative feedback condition so that positive effects would indicate that IM was higher in the negative feedback condition, and negative effects would indicate that IM was lower in the negative feedback condition relative to the comparison group. We also converted this mean difference to a Hedges'  $g$  with small sample bias correction (i.e., unbiased estimate of  $g$ ; Hedges and Olkin 1985, p. 80).

When possible, we calculated effect sizes from means, standard deviations, and sample sizes. When this information was not reported in a study, corresponding inference test statistics were used to derive an effect size. In the case when sample size information was unavailable, we used the inference test with assumed equal sample sizes (see Rosenthal 1994). If statistical significance was denoted yet both raw data and inferential test statistics were unavailable, a conservative effect size was derived with an assumed  $p$  value of 0.05 (Rosenthal and Rubin 2003). Although this approach may underestimate the effect, the alternative would be to exclude such data, which would overestimate the effect and aggravate the file drawer problem.

This occurred for 10 out of 78 included studies.<sup>1</sup> These conversion formulae are consistent with meta-analytic methods reported by Cooper et al. (2009).

## Coder Reliability

All reports were coded independently by trained coders that included the first, third, and fourth authors. The coders had experience coding for meta-analysis and were extensively trained on each code using the previously mentioned coding frame. As a reliability check, all pairs of codes for each study were compared for agreement between the two coders. We calculated a reliability measure between coding by dividing the number of matched codes by the total number of codes. Half of the included studies were double-coded to establish reliability before coding the remaining studies. Problematic codes during reliability calculation were the effect size value and feedback characteristics, such as attributional and autonomy-supportive feedback. The reliabilities for the problematic codes were 82.8 and 80.3% for effect size value and feedback characteristics, respectively. Therefore, these two categories of codes were double-coded. The reliability between inter-raters without counting the problematic codes was 95.4%. If there were any disagreements, codes from a third coder were consulted and used for resolution.

## Data Integration

Before conducting any meta-analytic procedures, we counted the number of positive and negative effects. Then, we calculated the range of estimated relationships of negative feedback on motivation and related outcomes. In addition, we examined the distribution of effect sizes and sample sizes to inspect for any statistical outliers by applying Grubbs' (1950) test, which identifies outliers beyond three standard deviations. Due to the possibility of not obtaining all relevant studies, we employed Duval and Tweedie's (2000) trim-and-fill procedure to assess if effect sizes differed from normally distributed estimates under random effects. We also ran funnel plots to help judge the impact of publication bias.

**Calculating Average Effect Sizes** A weighting procedure was used to calculate average effect sizes across independent samples. Each effect size was first multiplied by the inverse of its variance; then, the sum of these products was divided by the sum of their inverses. This procedure allows more weight to samples of larger size, which is generally preferred (Hedges and Olkin 1985) since larger samples give more precise population estimates. In addition, we calculated 95% confidence intervals for weighted average effect sizes; if the interval did not contain zero, the null hypothesis that negative feedback had no effect on motivation was rejected.

<sup>1</sup> We conducted moderator analyses comparing effects from studies using assumed information to derive effect sizes (ES) with studies with studies using more precise algebraic formulae. For the negative feedback condition vs. neutral/no feedback conditions, there was no evidence of significant moderation of effect size treatment (FE:  $Q = .27, p = .60$ , RE:  $Q = 1.16, p = .28$ ). Regardless of effect size treatment, both weighted average effects ( $k = 13$  treated effect sizes;  $k = 32$  non-treated effect sizes) were not significantly different from zero. For the negative feedback condition vs. positive feedback conditions, there was no evidence of significant moderation of ES treatment for random effects (RE:  $Q = 1.97, p = .16$ ), but there was evidence of moderation under fixed effects (FE:  $Q = 15.81, p = .001$ ). However, under fixed effects, effects for both treated ES ( $k = 15$ ) and non-treated ES ( $k = 85$ ) were both in the same direction and significantly different from zero, which supports the main effect that was found. Although non-treated ES studies had larger effects, this is expected because the treated effects were conservative estimates. Also, it should be noted that the sample sizes in the subgroup analysis were heavily unbalanced, and results should be interpreted with some caution as well.

**Identifying Independent Hypothesis Tests** When calculating effect sizes, determining whether an effect size is independent (participants in one sample providing the observations do not overlap with another sample) can be problematic when there are multiple effect sizes from a single sample (i.e., multiple levels of potential moderators). Therefore, we used a shifting unit of analysis approach (Cooper 1998) as a method for dealing with dependent effect sizes. This approach involved coding as many effect sizes from each study as exist as a result of variations in characteristics of the manipulation, sample, and outcomes within the study. When calculating the overall effect size, multiple effect sizes were averaged to create a single effect size for each sample. However, when examining moderators, we allowed a single sample to contribute an effect size to each level of the moderator.

**Heterogeneity and Moderator Analysis** Effect sizes may vary even if they estimate the same underlying population value. Therefore, homogeneity analyses were needed to determine whether the features of studies contributed to this variance above and beyond sampling error alone. We tested homogeneity of the observed set of effect sizes using a within-class goodness-of-fit statistic ( $Q_w$ ),  $I^2$ , and  $Tau^2$ . A significant  $Q_w$  statistic suggests that sampling variation alone could not adequately explain the variability in the effect size estimation and it follows that moderator variables should be examined (Cooper et al. 2009). Similarly, homogeneity analyses can be used to determine whether multiple groups of average effect sizes vary more than predicted by sampling error. In this case, statistical differences among different categories of studies were tested by computing the between-class goodness-of-fit statistic,  $Q_b$ . A significant  $Q_b$  statistic indicates that average effect sizes vary between categories of the moderator variables more than predicted by sampling error alone.  $I^2$  indicates the proportion of the observed variance reflecting differences in true effect sizes, rather than sampling error.  $Tau^2$  is the variance of the true effect sizes. In addition, we also calculated 95% prediction intervals (Borenstein et al. 2017), which indicates the range the effect size will fall in a given population drawing from the entire universe of studies.

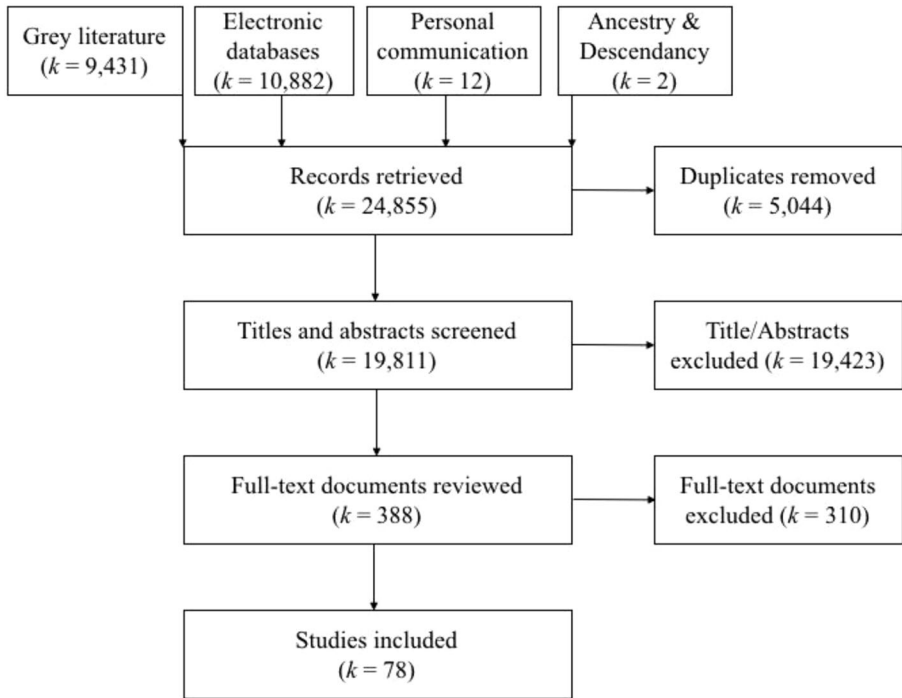
For continuous moderators such as percentage of females or age in the sample, we used meta-regression to assess moderation of these demographic characteristics. We also used two estimation procedures, method of moments and full maximum likelihood, when modeling fixed and random effects, respectively. A sample meta-regression equation was as follows:

$$ES_i = \beta_0 + \beta_{\text{moderator}} + u_i + e_i$$

In general, the criteria we employed to determine if a moderating factor can be meaningfully analyzed was if there were at least four unique study samples per level/group in a given comparison. We argue that below this number, moderator tests would be underpowered. In addition, we determined that if 80% or more of the studies did not report information about the moderator, a test of moderation would be invalid. Because we would be assuming a particular characteristic for a large majority of the studies, this would misrepresent how relevant details were simply unreported. Thus, a number of theoretical moderators could not be examined such as ethnicity, race, country of origin, and evaluator and feedback receiver characteristics.

**Fixed and Random Effects** In a fixed effects (FE) model of error, we assumed that the only source of error explaining why the effect size varies from one study to the next is sampling error or differences among participants across studies. In a random effects (RE) model of error, a study-level variance component is assumed to be an additional source of random variation. Due





**Fig. 1** PRISMA flow diagram. Note: Hand searches did not yield any unique studies after our search methods were employed

to the potential to over- or under-estimate error variance in moderator analyses (Hedges and Vevea 1998), we conducted all the analyses twice using both models of error. As a form of sensitivity analyses in order to examine the effect of different assumptions on the outcomes, we explored how different models changed the meta-analytic results (Greenhouse and Iyengar 1994). In the case of divergent results, effects were interpreted with these discrepancies in mind; in general, statistical significance under both models was considered more robust (Cooper 1998).

## Results

After title and abstract screening from our array of search strategies discussed in the method section, we identified 388 potentially relevant reports. Screening the full-text documents resulted in 78 included studies that examined the effect of negative feedback on IM compared to positive, neutral, no feedback, or an additional form of negative feedback. The primary reasons for excluding studies were a lack of feedback valence manipulation, no negative feedback comparisons, and no IM outcomes. See Fig. 1 for PRISMA flow diagram (Moher et al. and The PRISMA Group 2009) of the retrieval process. The included studies appeared between the years 1972 and 2016. The sample sizes ranged from 8 to 358. Participants of the included studies ranged from primary/secondary school students to postsecondary students. All participants received bogus feedback in laboratory contexts (Table 2).

First, we examined whether the three comparisons of negative feedback versus positive feedback ( $k = 100$ ), negative feedback versus neutral feedback ( $k = 10$ ), and negative

**Table 2** Characteristics of included studies

	<i>k</i>	Percent
Publication type		
Peer-reviewed journal article	51	65.4
Doctoral dissertation	19	24.4
Master's thesis	5	6.4
Conference paper	3	3.8
Publication year		
1970s	8	10.3
1980s	23	29.5
1990s	19	24.3
2000s	17	21.8
2010–2016	11	14.1
Region		
USA	59.5	76.2
Non-USA	17.5	22.4
Not reported	1	1.3
Grade		
K-12	21	26.9
Postsecondary	57	73.1

Shu and Lam (2016) used samples from China and the USA; therefore, half-counts were given for *k* under region

feedback versus no feedback ( $k = 35$ ) were all significantly different from one another. The average effect of negative feedback compared to positive feedback was significantly larger when compared to neutral feedback (FE:  $Q(1) = 12.21, p < .001$ , RE:  $Q(1) = 1.60, p = .21$ ) and compared to no feedback (FE:  $Q(1) = 65.77, p < .001$ , RE:  $Q(1) = 7.74, p < .01$ ). There was no significant difference between the comparison with neutral feedback and the comparison with no feedback (FE:  $Q(1) = 1.59, p = .21$ , RE:  $Q(1) = 0.02, p = .90$ ). Therefore, these two categories were collapsed into a single category we called the *neutral or no feedback* condition. There is also a conceptual reason to combine these two conditions. Namely, in contrast to positive and negative feedback which convey an evaluative judgment, neutral feedback is essentially a non-evaluative statement (i.e., “Let’s move on to the next one”; Pretty and Seligman 1984, p. 1244). Likewise, there is no evaluation provided in the absence of feedback. In other words, negative and positive feedback contain evaluative information, whereas both the absence of feedback and neutral feedback do not.

From the included studies, we extracted 431 separate effect sizes based on 102 separate samples. The authors, sample sizes, and averaged effects for these studies along with other important study characteristics are listed in Table S1 in the online supplemental material. Of the effect sizes for IM, 311 represented the effect of negative feedback compared to positive feedback, 100 of which were overall effects collapsed across subgroups. Eighty-two effect sizes represented the effect of negative feedback on IM compared to neutral or no feedback, 45 of which were overall effects collapsed across subgroups. For the comparisons of negative feedback conditions only, there were 38 effect sizes extracted from 19 studies, 24 of which were overall effects collapsed across subgroups. Of the effect sizes for perceived competence, 41 represented the effect of negative feedback compared to positive feedback, 36 of which were overall effects collapsed across subgroups. Nine effect sizes represented the effect of negative feedback on IM compared to neutral or no feedback, eight of which were overall effects collapsed across subgroups. For the comparisons of negative feedback conditions only, there were six effect sizes, all of which were overall effects.

**Table 3** Results of analyses examining the overall effect of negative feedback compared to neutral or no feedback and positive feedback

	<i>k</i>	<i>g</i>	95% CI	95% PI	$Q_i/Tau^2$	<i>I</i> <sup>2</sup>	Number of effects			Range
							Pos	Zero	Neg	
<b>Intrinsic Motivation</b>										
Negative vs. Neutral or No Feedback	45	0.07 (-0.06)	-0.01, 0.14 (-0.26, 0.15)	-1.32, 1.20	297.86 ***/0.38	85.23	24	4	17	-1.74, 1.43
Negative vs. Positive Feedback	100	-0.37*** (-0.36)***	-0.42, -0.32 (-0.48/-0.24)	-1.42, 0.68	491.41 ***/0.28	79.85	23	6	71	-2.91, 1.97
<b>Perceived Competence</b>										
Negative vs. Neutral or No Feedback	8	-0.48*** (-0.51)	-0.66, -0.31 (-1.07, 0.05)	-2.45, 1.43	61.41 ***/0.55	88.60	2	1	5	-1.58, 1.36
Negative vs. Positive Feedback	36	-0.90*** (-1.00)***	-0.98, -0.81 (-1.25, -0.75)	-2.34, 0.34	242.36 ***/0.42	86.80	1	0	35	-2.49, 0.26

95% CI confidence interval, 95% PI prediction interval, fixed-effects values are presented outside of parentheses and random-effects values are within parentheses. \**p* < .05; \*\**p* < .01; \*\*\**p* < .001. *I*<sup>2</sup> can be interpreted as representing considerable heterogeneity when over 75%. For the negative vs. positive comparison, trim-and-fill procedures detected five missing studies to the left of the mean using fixed effects only with intrinsic motivation. Imputed effects altered the weighted average effect size to *g* = -0.41 (95% CI -0.45, -0.36) under FE and *g* = -0.43 (95% CI -0.55, -0.30) under RE. For perceived competence, there were three effects missing from the left size of the distribution. Imputed effects were *g* = -1.09 (95% CI -1.17, -1.01) under FE and *g* = -1.11 (95% CI -1.38, -0.84) under RE

No effect size outliers were detected; however, one sample size outlier (Grouzet et al. 2004) with a large sample size was detected and Winsorized to its nearest neighbor. In the subsequent sections, we presented our analyses in two sets: (a) negative versus neutral or no feedback and (b) negative versus positive feedback.

### Overall Effects of Negative Feedback Compared to Neutral or No Feedback

Acknowledging that negative feedback is the active agent in the comparison with neutral or no feedback, we present these results first. When comparing with positive feedback, we introduce a secondary influence that may contaminate the effect, making it more difficult to disentangle the impact on IM due to a joint effect of the feedback valences. Table 3 presents the descriptive information regarding effect sizes and captures the overall effect of negative feedback on IM and perceived competence. Overall, under fixed and random effects, there was no significant difference between negative feedback and neutral feedback on IM ( $k=45$ ). Main effects were  $g=.07$  under fixed effects and  $g=-0.06$  under random effects, and both were not significantly different from zero. For perceived competence ( $k=8$ ), negative feedback had a significant negative effect when compared to neutral feedback ( $g=-0.48$ ) under fixed effects, but not under random effects ( $g=-0.51$ ). Trim-and-fill procedures revealed no missing effect sizes in the distribution. A funnel plots for this analysis is presented in the online supplementary materials (Figs. 1 and 2); inspecting the plots did not reveal substantial publication bias.

In order to further understand the direction of these effects, we conducted an additional analysis using only studies that provided measures of IM before and after receiving feedback (three studies: Lim 2005; Richards 1991; Viciania et al. 2007;  $k=7$  samples). We calculated appropriate effect sizes (Morris and DeShon 2002) and analyzed the change in IM from pre- to post-test by each condition (negative feedback, neutral/no feedback). These results demonstrated that on average, there was a decline in IM following negative feedback (FE:  $g=-0.41$ , 95% CI= $-0.62, -0.21$ ; RE:  $g=-0.44$ , 95% CI= $-0.81, -0.07$ ). For neutral/no feedback, there was no significant difference in IM from pre- to post-test (FE:  $g=0.13$ , 95% CI= $-0.09, 0.34$ ; RE:  $g=0.13$ , 95% CI= $-0.17, 0.42$ ). The change of IM for negative feedback was significantly different than the change of IM for neutral/no feedback (FE:  $Q=12.61$ ,  $p<.001$ ; RE:  $Q=5.47$ ,  $p<.05$ ). Although based on a handful of studies that may not necessarily be representative of all included studies, these findings indicate a decrease of IM following negative feedback and a slight positive (not significantly different from zero) effect following neutral/no feedback. This pattern of results was mostly aligned with our findings using the larger group of included studies that did not provide pre-test IM measures.

### Moderator Analyses for Negative Feedback Compared to Neutral or No Feedback

Due to sufficient heterogeneity among effect sizes in the main effects analysis, we assessed moderators of the effect of negative feedback compared to neutral or no feedback, but only for the IM outcome. We did not run moderator analyses for the perceived competence outcome because small samples sizes and limited variability restricted our ability to draw meaningful comparisons. We examined eight moderators related to publication status,

**Table 4** Results of moderator analyses for negative feedback compared to no neutral feedback on IM

Outcome/Moderator	<i>k</i>	<i>g</i>	95% CI	95% PI	<i>Q<sub>b</sub></i>	<i>Q<sub>w</sub></i>	<i>I<sup>2</sup></i>	<i>Tau<sup>2</sup></i>
Publication type								
Published	15	0.41*** (0.23)	0.29, 0.52 (-0.08, 0.54)	-1.31, 1.76	54.62*** (5.47)*	251.25***	88.24	0.48
Unpublished	30	-0.16*** (-0.24)	-0.26, -0.07 (-0.48, 0.004)	-0.75, 1.23			76.57	0.22
Social context of delivery								
In person	37	0.17*** (0.01)	0.09, 0.25 (-0.20, 0.23)	-1.25, 1.27	39.81*** (4.15)*	265.62***	84.54	0.37
Mediated	8	-0.46*** (-0.49)*	-0.65, -0.29 (-1.00, -0.09)	-1.56, 0.58			74.69	0.15
Feedback standard								
Criterion	24	0.18*** (-0.09)	0.00, 0.27 (-0.36, 0.18)	-1.51, 1.33	17.91*** (0.08)	287.96***	89.13	0.45
Normative	21	-0.16* (-0.03)	-0.28, -0.03 (-0.33, 0.28)	-1.13, 1.07			73.84	0.25
Motivation features								
Standard	28	-0.01 (0.05)	-0.13, 0.10 (-0.21, 0.28)	-1.05, 1.15	28.82*** (1.86)	243.65***	75.51	0.27
Instructional	6	0.30*** (0.07)	0.13, 0.47 (-0.53, 0.68)	-2.08, 2.22			91.78	0.50
Effort	2	-0.30 (0.11)	-0.72, 0.13 (-1.40, 1.62)	-			87.89	0.99
Autonomy-supportive	5	-0.26 (-0.43)	-0.58, 0.06 (-1.05, 0.18)	-2.44, 1.57			68.86	0.30
Praise	3	0.54*** (0.06)	0.29, 0.80 (-1.36, 1.57)	-18.19, 18.31			96.11	1.47
Task Interestingness								
Interesting	18	0.09 (-0.12)	-0.06, 0.23 (-0.47, 0.22)	-1.83, 1.59	0.14 (0.18)	305.73***	85.96	0.62
Not reported	27	0.05 (-0.03)	-0.03, 0.14 (-0.29, 0.23)	-1.23, 1.17			85.92	0.32
Age								
K-12 students	18	0.30*** (0.07)	0.19, 0.42 (-0.23, .36)	-1.40, 1.54	29.67*** (1.28)	276.19***	86.42	0.46
College students	27	-0.11* (-0.17)	-0.21, -0.02 (-0.43, .10)	-1.15, 1.09			82.16	0.28
Gender								
Male	8	0.08 (0.01)	-0.18, 0.35 (-0.55, 0.58)	-1.76, 1.78	0.19 (0.06)	31.23***	72.59	0.44
Female	6	-0.01 (0.03)	-0.32, 0.30 (-0.34, 0.40)	-0.68, 0.74			12.05	0.03
Measurement								
Self-report	22	0.05 (-0.18)	-0.05, 0.15 (-0.47, 0.11)	-1.81, 1.45	0.18 (1.45)	309.01***	91.59	0.59
Behavior	24	0.08 (0.08)	-0.03, 0.20 (-0.22, 0.37)	-0.75, 0.91			61.22	0.14

Fixed-effects values are presented outside of parentheses and random-effects values are within parentheses. Superscript letters denote significant pairwise comparisons. \**p* < .05; \*\**p* < .01; \*\*\**p* < .001

feedback characteristics, task characteristics, sample characteristics, and methodological characteristics (see Table 4).<sup>2</sup>

## Publication Status

First, we examined the association between the magnitude of effect sizes on IM and the publication status of the study report. Effects from published reports were significantly more positive than those from unpublished sources under a fixed-error model, but not under a random-error model.

## Feedback Characteristics

Feedback characteristics that we examined as moderators included the social context of feedback delivery, reference to a criterion versus normative standard, and motivational features of the feedback.

**Social Context of Feedback Delivery** After excluding one study (Butler 1998) due to lack of reporting, the effect of negative feedback delivered in-person was positive (though not significant under random effects assumptions). The effect of negative feedback in mediated forms was negative. This difference between in-person and mediated feedback was significant under both fixed-error and random-error assumptions.

**Normative Versus Criterion Standard** We excluded one study (Vallerand and Reid 1984) that did not provide information on the feedback standard. Under fixed-effects assumptions only, normative feedback had a negative effect on IM, whereas criterion-based feedback had a positive effect. The average weighted effect of negative feedback compared to neutral or no feedback significantly varied depending on whether feedback was normative or criterion-based. This difference was significant under fixed-error assumptions only.

**Feedback with Motivation Features** Next, we assessed whether there were differences between feedback that included motivational features such as attributions toward ability or effort, controlling or autonomy-supportive language, praise, and instructional details. Some studies combined multiple motivational features such as instruction and controlling language (e.g., Lim 2005). The average weighted effect of negative feedback compared to neutral or no feedback significantly varied across feedback with various motivation features under fixed-error assumptions only. Under fixed-effects assumptions, there were also several significant pairwise comparisons between effect size categories based on motivational features. The effects of instructional negative feedback ( $Q(1) = 9.20, p < .001$ ) and negative feedback that included praise ( $Q(1) = 15.08, p < .001$ ) were both positive and significantly different than effects of standard feedback with no motivation features. The effect of negative feedback with praise also significantly differed from that of autonomy-supportive feedback ( $Q(1) = 14.58, p < .001$ ) and from effort feedback ( $Q(1) = 10.98, p < .001$ ). The effect of instructional feedback was

<sup>2</sup> Because we tested each moderator separately, there is a possibility that moderators were confounded with one another. Therefore, we examined pairwise relationships between significant moderator variables: social context of feedback, feedback standard, motivational feedback types, task interestingness, and participant age group, and measurement type. Because all variables assessed were categorical, chi-square tests were conducted; however, none of the results were significant, suggesting no evidence of confounding moderators.

significantly more positive than effort feedback ( $Q(1) = 6.55, p < .01$ ) and autonomy-supportive feedback ( $Q(1) = 9.12, p < .001$ ). There were no other significant pairwise comparisons.

## Task Characteristics

Next, we examined whether negative feedback effects differed depending on level of task interestingness. The effect of negative feedback was not significantly different for tasks that were identified as interesting compared to tasks that did not report on it.

## Sample Characteristics

For sample characteristics, we examined whether negative feedback effects varied depending on the gender and age of the feedback receiver. In regard to participants' gender, group comparison moderator analyses found no evidence of significant differences between effects for men and women. Meta-regression results using percent female in the sample composition ( $k = 36$ ) suggested that gender was not a significant moderator under fixed effects ( $\beta = -0.07, p = .67; Q_{\text{model}} = .13, p = .71; Q_{\text{residual}} = 29.18, p = .70$ ) or random effects ( $\beta = -0.15, p = .71; Q_{\text{model}} = .19, p = .67; Q_{\text{residual}} = 286.61, p < .001$ ). Regarding participant age, under fixed-error assumptions only, the effect of negative feedback compared to neutral or no feedback was significantly different for college students than younger students. When comparing negative with neutral or no feedback, studies with preschool to high school students showed increases in IM, and studies with college students had decreases in IM. We also examined age as a moderator by meta-regressing the mean age of the sample (when the relevant information was reported,  $k = 38$ ) on the effect sizes. Under fixed effects, the slope value for age on effect size was negative and significantly different from zero (FE:  $\beta = -0.06, p < .001; Q_{\text{model}} = 51.60, p < .001; Q_{\text{residual}} = 235.20, p < .001$ ), but not under random effects (RE:  $\beta = -0.04, p = .10; Q_{\text{model}} = 2.64, p = .10; Q_{\text{residual}} = 31.75, p = .58$ ). This suggests that as age increases, the effect of negative feedback compared to no feedback is more detrimental to IM (i.e., a positive effect on IM that is reversing to a negative effect as age increases).

## Methodological Characteristics

The last set of moderators examined whether effects varied depending on methodological characteristics. In particular, we tested whether the outcome measurement type, self-reported or behavioral, was a moderator when comparing negative with neutral or no feedback. Under fixed and random-error assumptions, there was not a significant difference between behavioral measures and self-reported measures of IM.

## Overall Effects of Negative Feedback Compared to Positive Feedback

Table 3 presents descriptive information regarding effect sizes for the negative-positive feedback comparison and the overall effect of negative feedback compared to positive feedback for IM and perceived competence. Overall, IM was significantly lower after receiving negative feedback compared with after receiving positive feedback under both fixed and

Table 5 Results of moderator analyses for negative feedback compared to positive feedback on IM

Outcome/Moderator	<i>k</i>	<i>g</i>	95% CI	95% PI	$Q_b$	$Q_w$	$I^2$	$Tau^2$
Publication type								
Published	76	-0.32*** (-0.30)***	-0.38, -0.26 (-0.43, -0.16)	-1.29, 0.69	10.36** (3.35)†	481.05***	77.34	0.24
Unpublished	24	-0.52*** (-0.58)***	-0.63, -0.42 (-0.82, -0.34)	-1.90, 0.74			84.68	0.39
Social context of delivery								
In person	81	-0.36*** (-0.36)***	-0.42, -0.30 (-0.49, -0.22)	-1.24, 0.52	0.72 (0.05)	490.70***	72.71	0.19
Mediated	19	-0.42*** (-0.39)*	-0.53, -0.30 (-0.66, -0.12)	-2.03, 1.25			90.89	0.43
Feedback standard								
Criterion	46	-0.31*** (-0.38)***	-0.39, -0.24 (-0.55, -0.21)	-1.40, 0.64	4.15* (0.06)	486.79***	77.26	0.25
Normative	54	-0.42*** (-0.35)***	-0.49, -0.35 (-0.52, -0.18)	-1.49, 0.79			81.66	0.31
Motivation features								
Standard	68	-0.37*** (-0.38)***	-0.43, -0.32 (-0.51, -0.25)	-1.28, 0.53	64.72*** (4.01)	467.22***	82.36	0.20
Instructional	6	0.23** (-0.23)	0.07, 0.39 (-0.88, 0.41)	-2.53, 2.07			92.72	0.34
Ability	5	-0.58*** (-0.51)*	-0.80, -0.36 (-1.02, -0.01)	-2.32, 1.30			80.04	0.26
Effort	5	-0.59*** (-0.59)***	-0.84, -0.33 (-0.84, -0.33)	-1.01, -0.17			0	0
Autonomy-supportive	4	-0.61** (-0.87)*	-0.99, -0.23 (-1.64, -0.09)	-4.06, 2.32			70.45	0.39
Controlling	2	-0.68*** (-0.70)	-1.08, -0.28 (-1.83, 0.44)	-			87.59	0.34
Task Interestingness								
Interesting	29	-0.27*** (-0.20)†	-0.36, -0.18 (-0.42, 0.02)	-1.35, 0.94	6.82** (3.04)†	484.45***	81.83	0.30
Not reported	71	-0.42*** (-0.43)***	-0.48, -0.35 (-0.58, -0.29)	-1.48, 0.62			78.82	0.27
Age								
K-12 students	32	-0.42*** (-0.34)**	-0.52, -0.33 (-0.55, -0.13)	-1.53, 0.85	1.78 (0.06)	489.63***	82.01	0.33
College students	68	-0.35*** (-0.37)***	-0.41, -0.28 (-0.52, -0.23)	-1.40, 0.66			78.88	0.26
Gender								
Male	11	-0.56*** (-0.62)***	-0.70, -0.42 (-0.86, -0.39)	-1.48, 0.24	0.05 (0.00)	41.37***	68.81	0.13
Female	7	-0.59*** (-0.62)***	-0.75, -0.43 (-0.92, -0.33)	-1.21, -0.03			35.52	0.03
Measurement								
Self-report	52	-0.43*** (-0.43)***	-0.50, -0.36 (-0.59, -0.26)	-1.92, 1.06	7.31** (1.18)	501.59***	82.46	0.29
Behavior	50	-0.28*** (-0.29)**	-0.37, -0.20 (-0.48, -0.12)	-1.78, 1.20			76.76	0.29

Fixed-effects values are presented outside of parentheses and random-effects values are within parentheses. 95% PI could only be calculated for analyses with three or more studies. \* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$ ; † $p < .10$



random effects. Main effects were  $g = -0.37$  under fixed effects and  $g = -0.36$  under random effects, and both were significantly different from zero. Trim-and-fill procedures detected five missing studies to the left of the mean using fixed effects only. Imputed effects altered the weighted average effect size to  $g = -0.41$  (95% CI =  $-0.45, -0.36$ ) under FE and  $g = -0.43$  (95% CI =  $-0.55, -0.30$ ) under RE. For perceived competence ( $k = 36$ ), negative feedback had a significant, large negative effect when compared to positive feedback under fixed effects ( $g = -0.90$ ) and random effects ( $g = -1.00$ ). Trim-and-fill procedures detected three missing studies to the left of the mean using fixed effects only. Imputed effects slightly increased the weighted average effect size to  $g = -1.09$  (95% CI =  $-1.17, -1.01$ ) under FE and  $g = -1.11$  (95% CI =  $-1.38, -0.84$ ) under RE. The funnel plots did not reveal substantial publication bias (presented in Figs. 3 and 4 in online supplementary materials).

We also examined studies that provided measures of IM before and after receiving feedback (four studies: Lim 2005; Richards 1991; Viciano et al. 2007; Weidinger et al. 2017;  $k = 8$  samples). Looking at the change in IM from pre- to post-test by each condition (negative feedback, positive feedback), we found a decline in IM following negative feedback (FE:  $g = -0.34$ , 95% CI =  $-0.51, -0.17$ ; RE:  $g = -0.40$ , 95% CI =  $-0.70, -0.09$ ). For positive feedback, there was a significant increase in IM from pre- to post-test (FE:  $g = 0.28$ , 95% CI =  $0.10, 0.09$ ; RE:  $g = 0.41$ , 95% CI =  $0.45, 0.74$ ). The change of IM for negative feedback was significantly different than the change of IM for positive feedback (FE:  $Q = 25.18$ ,  $p < .001$ ; RE:  $Q = 12.86$ ,  $p < .001$ ). These findings indicate a decrease of IM following negative feedback and a positive effect following positive feedback. This pattern of results was consistent with our prior findings, but we caution once more that these studies may not be representative of the larger pool of studies.

### Moderator Analyses for Negative Feedback Compared to Positive Feedback

Next, we assessed moderators of the effect of negative feedback compared to positive feedback on IM and perceived competence to explore the significant degree of heterogeneity present in the distribution of effect sizes. For moderator analyses of the effect of negative versus positive feedback, we examined the same feedback, task, sample, and methodological characteristics as in the previous moderator analyses comparing negative to neutral or no feedback, with the addition of controlling feedback as a motivational feature (Tables 5 and 6).

#### Publication Status

We found significant moderation of publication status for both intrinsic motivation and perceived competence outcomes. Effects from unpublished reports were significantly larger (more negative) than those from unpublished sources under a fixed-error model, but not under a random-error model (for intrinsic motivation, the RE model was marginally significant).

#### Feedback Characteristics

Next, we examined whether the following feedback characteristics explained variability in the effects: social context of feedback delivery, reference to a criterion versus normative standard, and motivational features. These characteristics applied to both positive and negative feedback conditions.

**Table 6** Results of moderator analyses for negative feedback compared to positive feedback on perceived competence

Outcome/Moderator	<i>k</i>	<i>g</i>	95% CI	95% PI	$Q_b$	$Q_v$	$I^2$	$Tau^2$
Publication type								
Published	23	-0.81*** (-0.86)***	-0.91, -0.71 (-1.21, -0.53)	-2.30, 0.58	9.44** (2.10)	232.55**	89.16	0.45
Unpublished	13	-1.15*** (-1.23)***	-1.28, -0.95 (-1.61, -0.85)	-2.58, 0.12			79.05	0.34
Social context of delivery								
In person	24	-0.90*** (-1.08)***	-1.01, -0.77 (-1.40, -0.77)	-2.38, 0.22	0.00 (1.03)	241.63***	82.50	0.37
Mediated	11	-0.90*** (-0.79)**	-1.02, -0.76 (-1.25, -0.34)	-2.60, 1.02			92.60	0.59
Feedback standard								
Criterion	12	-0.67*** (-1.07)***	-0.84, -0.50 (-1.59, -0.55)	-2.91, 0.77	9.52** (0.05)	201.45***	87.51	0.61
Normative	20	-0.96*** (-0.94)***	-1.07, -0.86 (-1.22, -0.66)	-2.09, 0.21			83.86	0.28
Motivation features								
Standard	28	-0.96*** (-0.97)***	-1.06, -0.86 (-1.20, -0.74)	-2.01, 0.07	15.41*** (0.71)	209.20***	78.27	0.24
Instructional	2	-0.40* (-0.58)	-0.75, -0.05 (-1.50, 0.34)	-			81.25	0.35
Ability	5	-0.63*** (-1.09)*	-0.84, -0.43 (-2.13, -0.06)	-5.07, 2.89			95.72	1.29
Task Interestingness								
Interesting	14	-1.41*** (-1.33)***	-1.54, -1.27 (-1.72, -0.94)	-2.67, 0.01	88.70*** (8.64)**	153.12***	87.24	0.34
Not reported	22	-0.56*** (-0.74)***	-0.67, -0.45 (-0.97, -0.51)	-1.68, 0.20			73.25	0.19
Age								
K-12 students	4	-1.86*** (-1.90)***	-2.19, -1.53 (-2.78, -1.01)	-5.93, 2.13	35.44*** (6.98)**	206.20***	85.59	0.67
College students	29	-0.82*** (-0.88)***	-0.91, -0.74 (-1.26, -0.64)	-2.12, 0.36			85.44	0.35
Gender								
Male	3	-1.58*** (-1.82)***	-1.94, -1.22 (-2.82, -0.82)	-18.68, 15.04	2.28 (0.63)	36.63***	93.09	1.50
Female	6	-1.17*** (-1.25)***	-1.51, -0.84 (-1.99, -0.52)	-2.61, 0.11			34.99	0.10

Fixed-effects values are presented outside of parentheses and random-effects values are within parentheses; 95% PI could only be calculated for analyses with three or more studies.

\* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$

**Social Context of Feedback Delivery** After excluding two studies (Butler 1998; Tang 1990) for which the information provided was insufficient to determine the social context of feedback delivery, there was no difference in the effect of negative feedback when delivered in-person compared to when feedback was mediated in other modes. For perceived competence, there was no evidence of a significant difference between the two feedback delivery groups.

**Normative Versus Criterion Standard** When assessing the difference between normative and criterion-based feedback, we excluded four studies because they did not specify the feedback standard (Tang and Sarsfield-Baldwin 1991; Vallerand and Reid 1984; Viciano 2007; Woodcock 1990). Two studies (Dyck et al. 1979; Burgers et al. 2015) included both kinds and contributed two effect sizes each. The average weighted effect of negative feedback compared to positive feedback varied depending on whether feedback was normative or criterion-based under fixed-error assumptions only. Normative negative feedback compared to positive feedback had a significantly larger negative effect on IM.

For the perceived competence outcome, we excluded three samples because they did not specify the feedback standard (Vallerand and Reid 1984; Woodcock 1990, two samples). Feedback standard did not significantly explain variation in effects on perceived competence under either error model.

**Feedback with Motivation Features** Next, we assessed whether there were differences in the effect of negative feedback depending on whether other motivational features were included. The average weighted effect of negative feedback compared to positive feedback significantly varied across feedback with different types of motivation features under fixed-error assumptions only. We then proceeded to conduct pairwise comparisons under fixed-effects assumptions. The only positive effect was for instructional negative feedback, which was significantly different from the negative effects of standard feedback ( $Q(1) = 51.08, p < .001$ ), ability feedback ( $Q(1) = 4.01, p < .05$ ), effort feedback ( $Q(1) = 28.43, p < .001$ ), controlling feedback ( $Q(1) = 17.27, p < .001$ ), and autonomy-supportive feedback ( $Q(1) = 15.77, p < .001$ ).

For perceived competence, there were only two types of feedback with enough contributing effect sizes to conduct moderator analyses in addition to standard feedback without any motivation features: ability-focused feedback and instructional feedback. The effect of negative compared to positive feedback on perceived competence varied significantly across feedback with different motivational features under fixed-error but not random-error assumptions. The average effect of all types of negative compared to positive feedback on perceived competence was negative. However, pairwise comparisons under fixed-effects assumptions revealed that the average negative effect on perceived competence for standard feedback was greater than that for instructional feedback,  $Q(1) = 15.47, p < .001$ . Instructional feedback also had a smaller negative effect compared to ability feedback,  $Q(1) = 13.20, p < .001$ .

## Task Characteristics

Results revealed that the negative effect of negative feedback compared to positive feedback was significantly smaller for tasks that were identified as interesting compared to tasks that did not report on interestingness. This effect was significant under fixed-error assumptions and marginally significant under random-error assumptions. For perceived competence, this effect was reversed. Negative feedback had a significantly larger negative effect for interesting tasks compared to tasks that did not specify interestingness under both error assumptions.

## Sample Characteristics

Regarding the moderation of gender, using studies that reported effects comparing negative and positive feedback for one gender only, we compared effect sizes for men and women. For intrinsic motivation and perceived competence, there were no significant differences between male and female participants. Second, results from the meta-regression ( $k = 84$ ) indicated that the slope coefficient of female percentage was significant under fixed effects (FE:  $\beta = -0.16$ ,  $p = .04$ ;  $Q_{\text{model}} = 4.24$ ,  $p = .04$ ;  $Q_{\text{residual}} = 546.16$ ,  $p < .001$ ), but not under random effects (RE:  $\beta = -0.01$ ,  $p = 0.95$ ;  $Q_{\text{model}} = 0.005$ ,  $p = .95$ ;  $Q_{\text{residual}} = 91.43$ ,  $p = .22$ ). This suggests that the effect of negative compared to positive feedback on IM decreased as female percentage increased. For perceived competence ( $k = 27$ ), meta-regression results revealed no evidence of a significant effect of percent female of the sample under fixed ( $\beta = 0.04$ ,  $p = .84$ ;  $Q_{\text{model}} = 0.04$ ,  $p = .84$ ;  $Q_{\text{residual}} = 207.11$ ,  $p < .001$ ) or random effects ( $\beta = 0.45$ ,  $p = .40$ ;  $Q_{\text{model}} = 0.70$ ,  $p = .40$ ;  $Q_{\text{residual}} = 23.70$ ,  $p = .54$ ), suggesting that the effect of negative compared to positive feedback on perceived competence was unrelated to the percentage of females in the sample.

For age, under both fixed and random effects, there was no significant difference between college and preschool to twelfth grade students in the effect of negative compared the positive feedback on IM. Likewise, meta-regressing mean age (for studies that reported the relevant informant,  $k = 90$ ) on effect sizes suggested that the slope for age on effect size was not significantly different from zero under fixed effects ( $\beta = -0.003$ ,  $p = .42$ ;  $Q_{\text{model}} = 0.66$ ,  $p = .42$ ;  $Q_{\text{residual}} = 549.74$ ,  $p < .001$ ) or random effects ( $\beta = 0.005$ ,  $p = .69$ ;  $Q_{\text{model}} = 0.16$ ,  $p = .69$ ;  $Q_{\text{residual}} = 90.99$ ,  $p = .23$ ). For perceived competence, we found a significant difference in the effect of negative compared to positive feedback between the college student samples and preschool to twelfth grade samples under fixed-error assumptions and a marginally significant difference under random-error assumptions. There was a stronger negative effect for preschool to twelfth grade participants compared to college student participants. Similarly, meta-regression results ( $k = 23$ ) supported a developmental trend under fixed effects ( $\beta = 0.14$ ,  $p < .001$ ;  $Q_{\text{model}} = 17.97$ ,  $p < .001$ ;  $Q_{\text{residual}} = 150.95$ ,  $p < .001$ ) and random effects ( $\beta = 0.15$ ,  $p = .03$ ;  $Q_{\text{model}} = 4.86$ ,  $p = .03$ ;  $Q_{\text{residual}} = 16.38$ ,  $p = .75$ ), suggesting that as participant age increased, the effect of negative feedback compared to positive feedback on perceived competence became less negative.

A few studies compared the effects of negative feedback and positive feedback by ability and motivation levels of participants. Due to the small number of studies that examined these characteristics and the large amount of heterogeneity among those studies that did examine these characteristics, we did not conduct formal analyses on these variables. Superscripts in Table S1 indicate which studies presented results by these individual attributes.

## Methodological Characteristics

Under fixed effects only, there was a significant difference between types of IM measurements. Studies using behavioral measures of IM had smaller negative effects than studies using self-report measures. One effect (Anderson and Rodin 1989) was excluded from the moderator analysis because it combined behavioral and self-report measures as one index. This moderator analysis was not conducted on the perceived competence outcome, because all measures of perceived competence were self-reported.

## Analyses Comparing Different Forms of Negative Feedback: Feedback Characteristics

We were also interested in comparing the effects of different kinds of negative feedback. This comparison would elucidate what kind of negative feedback is the most and least likely to enhance IM and perceived competence. For these analyses, we limited the sample of studies to those that compared the effects of two or more different kinds of negative feedback. Due to a small number of studies that contributed effect sizes, we limited the comparisons to instructional versus non-instructional ( $k = 7$ ), effort-focused versus ability-focused ( $k = 3$ ), task-focused versus process-focused ( $k = 2$ ), non-threatening versus threatening ( $k = 6$ ), non-controlling versus controlling ( $k = 3$ ), and “wise” versus “unbuffered” ( $k = 4$ ). Table 7 presents results and examples of the various forms of negative feedback, and superscripts in Table S1 indicate which studies compared two forms of negative feedback.

For negative feedback comparisons, the average weighted effects on IM that were significantly different from zero under both fixed and random effects were only the comparisons between instructional versus non-instructional negative feedback and wise versus unbuffered negative feedback comparisons. Results favored instructional and wise negative feedback as more supportive of IM than their respective feedback counterparts. Only under fixed effects, the comparison of non-controlling feedback versus controlling feedback was also significantly different from zero. This highlighted that non-controlling negative feedback may also better support IM compared to negative feedback that used controlling language. The remaining feedback comparisons were not significantly different from zero under fixed and random effects.

For perceived competence, we looked at instructional versus non-instructional ( $k = 3$ ) and non-threatening versus threatening ( $k = 3$ ). For perceived competence, the average weighted effect of instructional negative feedback compared to non-instructional negative feedback was significantly different from zero under FE and RE assumptions, favoring instructional negative feedback. The effect of non-threatening compared to threatening negative feedback on perceived competence was not significant.

## Discussion

In the real world, people are regularly faced with providing performance evaluations that are often negative. Any expert in a wide variety of domains, perhaps on numerous occasions, will have to provide negative feedback. It is widely accepted that feedback can have a powerful influence to change behavior and improve learning. Despite the ubiquity of these commonly held beliefs, little empirical guidance has been available as to how negative feedback should be delivered for the greatest benefit for intrinsic motivation.

The central objective of this research synthesis was to measure the effect of negative feedback on intrinsic motivation. However, caution is needed when interpreting these findings because the overall effects mask heterogeneous effects of both negative and positive feedback. That is, when negative feedback is compared with no or neutral feedback, we can more confidently attribute effects to negative feedback rather than neutral feedback, which functions as a control condition. However, for the negative versus positive feedback comparison, it is more difficult to disentangle the effects of the negative feedback from positive feedback. In the following sections, we discuss these findings in further detail, highlighting trends that seemed consistent across conditions and models of error. We emphasize the findings that were significant under both fixed and random effects as most robust.

**Table 7** Summary of different forms of negative feedback and sample feedback statements

		<i>k</i>	<i>g</i>	95% CI		Example Feedback statements
				Low estimate	High estimate	
<i>Instructional</i> vs. non-instructional	IM	7	0.86 (0.95)	0.66 (.33)	1.07 (1.58)	<i>Instructional</i> : “However, one useful strategy for successful performance that is often overlooked in this puzzle is to search for words that run up diagonally, from left to right” <i>Non-instructional</i> : “You found (actual number) out of (8 more than they found). Your score indicates you did not do you very well on the puzzle” (Richards 1991).
	PC	3	0.56 (56)	0.10 (.10)	1.02 (1.02)	
Effort vs. ability	IM	7	0.14 (0.14)	−0.20 (−0.20)	0.49 (0.49)	<i>Ability</i> : “Compared to other children your age, you did not do so well. Remember it is important to solve more problems correctly than anyone else your age. When you solve more problems than other kids, it makes you look real smart. You look smarter when you get more problems correct.” <i>Effort</i> : “You must not have tried very hard to so poorly. Remember, the harder you try, the more you learn and the more problems you will solve. When you work very hard you get better at solving these problems” (Nichols et al. 1991).
Task vs. process	IM	2	0.02 (−0.01)	−0.35 (−0.60)	0.39 (0.58)	<i>Task</i> : “That’s not the right way to do it, because the blocks are not straightened up and are still messy.” <i>Process</i> : “The blocks are all crooked and in one big mess. Maybe you could think of another way to do it” (Kamins and Dweck 1999).
Non-threatening vs. threatening	IM	6	0.24 (0.32)	−0.01 (−0.21)	0.47 (0.85)	<i>Non-threatening</i> : “You did not do so well I’m kind of surprised, this is one of the lowest scores I have seen since I have been administering the test Is anything bothering you—did you not get enough sleep last night, or have you been working too hard lately. I think someone told
	PC	3	0.25 (0.42)	−0.20 (−0.45)	0.70 (1.29)	

**Table 7** (continued)

	<i>k</i>	<i>g</i>	95% CI		Example Feedback statements
			Low estimate	High estimate	
<i>Wise vs. unbuffered</i>	IM 4	0.44 (0.44)	0.03 (0.03)	0.84 (0.84)	me once that this particular test might be affected by things like that.” <i>Threatening</i> : You did quite badly I am surprised, this is one of the lowest scores I’ve seen since I’ve been administering the test, is something wrong” (Anderson and Rodin 1989) <i>Wise</i> : “I am giving you these comments because I have very high expectations and I know that you can reach them.” <i>Unbuffered</i> : “I am giving you these comments so that you’ll have feedback on your paper” (Yeager et al. 2014)
<i>Non-controlling vs. controlling</i>	IM 3	0.39 (0.26)	0.12 (–0.40)	0.66 (0.92)	<i>Non-controlling</i> : “You need to make more effort. Try to write a more original and interesting story rather than common story that people can expect easily. Read carefully to understand a given situation correctly and try to complete it. Even if a common thing comes up to your mind first, try to think about more original and flexible ideas.” <i>Controlling</i> : “Your grade results from comparing your story with other students’. After a little while, you SHOULD complete another tasks and your stories will be graded by comparing them with other students on the basis of the originality and interest.” (Lim 2005).

*IM* intrinsic motivation, *PC* perceived competence. Italicized moderators indicate the more beneficial moderator for difference effects that were significantly larger than zero. Values in parentheses indicate random effects; values outside parentheses indicate fixed effects

## Implications for Theory Development and Educational Practice

**Comparison with No or Neutral Feedback** Overall, results from 45 samples suggested that providing negative feedback seemed just as intrinsically motivating as giving no feedback, on average. In other words, there was no noticeable difference in IM as a function of receiving

negative feedback versus not receiving any feedback. Rather than benchmarking this effect to arbitrary cut-points of small, medium, and large, we interpreted this effect in relation to other relevant studies (Lipsey et al. 2012). The magnitude of negative versus neutral feedback was similar to what Kluger and Denisi (1996) found when examining the effect of praise feedback about the task on performance ( $d = .09$ ; Hattie and Timperley 2007). A recent meta-analysis with a similar outcome to our study found that the provision of rationale on autonomous motivation was  $d = .08$  and not significantly different from zero (Steingut et al. 2017). All in all, this is a small but typical effect in educational research.

This effect implies that what psychologists have classically thought of as the “mum effect” (Tesser and Rosen 1975), or shying away from providing any feedback, is not necessarily protecting the feedback receiver from a loss of IM. If anything, providing negative feedback allows the feedback receiver to obtain information about his or her progress, albeit negative. They can, in turn, use the information toward greater goal attainment. Because negative feedback may not necessarily reduce IM compared to neutral feedback, goal-setting theory asserts that all feedback, even negative, is beneficial for providing information. This resonates with what Deci et al. (1999) described as *informational* feedback. Understanding what the standard of excellence is and one’s progress in relation to this standard can prevent decreases in IM. This notion that negative feedback can be *better than no feedback at all* to the extent that it provides information that can facilitate future performance and goal attainment is also consistent with tenets of self-verification theory (Swann et al. 1988). Perhaps negative feedback is not perceived as demotivating compared to no feedback because one may desire negative feedback to verify self-views, even if those views are negative. It follows that feedback that maintains logical coherence between self-perceptions can maintain IM.

Although based on only eight study samples, compared to neutral/no feedback, negative feedback still moderately decreased one’s perceived competence. One possible explanation for the diminished sense of competence is that the effect on perceived competence served as a manipulation check. To assess whether negative feedback as a manipulation had its intended effect of communicating that a student had not successfully mastered some task, a student’s report of a loss of competence is expected, according to social cognitive theory (Bandura 1986) and SDT studies (Deci and Cascio 1972).

**Comparison with Positive Feedback** When compared with receiving positive feedback, negative feedback was detrimental to individuals’ IM on average based on 100 samples. The order of magnitude of this effect was similar to the main effect of choice on intrinsic motivation ( $d = .30$  to  $.36$ ; Patall et al. 2008). From Deci et al.’s (1999) meta-analytic review, their effect of rewards on intrinsic motivation is comparable to our effects of negative feedback compared to positive feedback ( $ds = -0.28$  to  $-0.40$ ). Altogether, this effect is a medium to large effect in educational research. This finding aligns with self-determination theory and its support for using verbal praise (but not expected, tangible, rewards as feedback) to increase IM (Deci et al. 1999). We observe a similar trend for perceived competence ( $k = 36$ ), but with a much larger effect, nearly three times the size.

**Main Effects Controlling for Pre-Measures of Intrinsic Motivation** Overall, the pattern of effects indicated (a) that the difference between negative and neutral feedback was negligible and (b) that positive feedback increased IM relative to negative feedback. Moreover, this finding is supported by our analyses when examining the change in IM before and after feedback, controlling for pre-measures of IM. Although based on just a handful of studies,



these findings indicate a decrease of IM following negative feedback, an increase of IM following positive feedback, and a slight positive (not significantly different from zero) effect following neutral/no feedback. One notable difference between analyses focused on pre-post changes and those focused on post-test only differences is that the pre-post effects indicate that on average, individuals did experience a decrease in intrinsic motivation that was significantly different from the negligible (positive) change in intrinsic motivation in the absence of evaluative feedback. On the other hand, post-test only results suggested that the average overall difference between negative and neutral or no feedback was negligible. However, it is important to note that the pre-post change effects are based on a very small number of studies that are not likely to be generalizable to the entire sample of studies included in this meta-analysis or the population of studies on this topic. Given the extensive heterogeneity in the nature of negative feedback and factors such as feedback instructiveness, criterion-based standards, and delivery mode that help to explain some of that heterogeneity, we assert that although negative feedback reduces intrinsic motivation under many conditions, various factors that mitigate this decline render the difference between negative and neutral or no feedback negligible on average. Clearly, more research examining pre- and post-test differences in IM following various forms of feedback is needed.

**Type of Feedback** Scholars and educators have debated what comprises the best kind of feedback to provide students. Results of this meta-analysis reveal important insights into this matter. First, across a number of analyses, instructional negative feedback whether compared to positive feedback ( $k = 6$ ), neutral feedback ( $k = 6$ ), or negative feedback without instructional information ( $k = 7$ ) had positive effects on IM. Constructive criticism or instructionally relevant feedback was found to be more intrinsically motivating than negative feedback that lacked such informational aspects (medium to large effect size differences). Because individuals seek to be seen in positive light, negative feedback without instructional details was on average detrimental to IM. Perhaps instructional feedback can provide the means to be competent in the future, thereby enhancing IM. Moreover, the powerful effect of including instructional strategies to accompany negative feedback supports tenets of goal theory. Constructive feedback can be the means through which a feedback receiver learns the next step toward a goal (Fong et al. 2016). Note that the responsibility of the student is critical when receiving feedback with directions to improve, playing an active role considering how to improve their performance in the future (Deci and Ryan 2016). This underscores the fact that negative feedback and its nuances are not always perceived as unpleasant.

Second, when comparing negative and neutral feedback, significant differences were found between criticism delivered in-person and feedback delivered in mediated forms. While in-person negative feedback enhanced IM relative to neutral feedback, negative feedback delivered through mediated forms diminished IM. Both were small effects based on 45 samples. This result was surprising because receiving criticism may be considered embarrassing, especially in the presence of others, as feedback is often delivered in front of an entire class. However, one could argue that the presence of another individual, the feedback giver, may impart a greater sense of care (Noddings 2002). The feedback receiver may interpret the negative feedback as less threatening when delivered in-person, which softens the blow of critical feedback. In-person feedback may communicate positive intentions to provide feedback for the receiver's benefit and improvement (Fong et al. 2018a). These in-person evaluative interactions may connote "problem-solving sessions" (Deci and Ryan 2016, p. 22), even though there may be no explicit dialogue. Furthermore, given that individuals tend to develop

a deeper connection with those who they view as instrumental for their goal pursuit (Fitzsimmons and Fishbach 2010), individuals may benefit from in-person criticism relative to no or neutral feedback. Perhaps because the human evaluator and his or her feedback is seen as useful toward goal attainment. Thus, in-person negative feedback provides an opportunity for both competence and relatedness support that is potentially missing from mediated feedback. Also, mediated feedback was mostly delivered using some sort of technology which has been found to have mixed responses from students (Winstone et al. 2017).

Third, when comparing negative feedback to neutral feedback, criterion-based negative feedback had positive effects on IM, whereas normative standards decreased IM (medium effect size difference from 45 samples). Compared to positive ( $k = 100$ ) feedback, negative feedback with criterion-based standards had a smaller negative effect than normative feedback. However, the difference between the effects was small and only under fixed effects. Altogether, criterion-based feedback is more intrinsically motivating; one possible reason is perhaps because normative comparisons promote more extrinsic forms of motivations that focus on competition, rather than encouraging intrinsic motivation. Our findings suggest that teachers should avoid feedback that compares students' performance with others. In light of the increasing national focus on standardized assessments to measure student learning and high-stakes testing, educators tend to be focused on normative comparisons of students' scores (Mangels et al. 2017). Results from our meta-analysis caution against implementing such practices. Rather our findings suggest that students may benefit from a greater focus on mastery and learning via criterion-based feedback.

It is also worth noting, though based on just a small sample of studies ( $k = 3$  to 4), that IM was greater when feedback included two particular attributes. The first is wise feedback which reinforces high expectations and provides personal assurance for students (Yeager et al. 2014). This is in contrast to overpraising mediocre work or to withholding negative feedback completely. These strategies often diminish students' IM, in spite of teachers intending to boost students' sense of self-esteem and trust. The second attribute is non-controlling language to bolster autonomy, and, in turn, IM. The use of controlling language instead causes individuals to resist feedback that attempts to control how they persist and engage in learning.

In sum, developing students' IM is one of the main educational objectives for instructors at all levels. The results of our meta-analysis indicated ways to improve student intrinsic motivation through specific feedback practices: mainly, providing feedback with instruction and directions to improve, without normative standards, and delivered in-person has the strongest impact on fostering student motivation. Additionally, we also suggest a more tentative recommendation to offer wise and autonomy-supportive feedback due to limited evidence. In order to aid teachers to deliver this kind of feedback, we encourage professional development efforts that enhance classroom discourse. Workshops and trainings that involve changing teachers' perspectives toward student learning and observations of their own teaching are effective ways to alter teachers' discourse practices, particularly increasing their use of constructive feedback (Kiemer et al. 2015).

**Task** The moderating influence of task characteristics was significant only when negative feedback was compared with positive feedback for IM ( $k = 100$ ), resulting in a small effect size difference. Because we did not find any significant moderator results when negative feedback was compared with neutral or no feedback, perhaps the task type moderation may be more attributed to the effect of positive feedback. When a task was interesting, negative feedback compared to positive feedback may not cause as great of a decline in IM because the receiver

was potentially still interested in the stimulating task. One could argue that the positive feedback may also be particularly impactful when a task is interesting. Doing well on a task that has a high degree of value may be more intrinsically motivating (Heyman and Dweck 1992). In addition, perhaps given the highly engaging nature of interesting tasks, it may be unlikely that negative feedback may hinder IM during an interesting task. Although negative feedback's effect on IM was negative overall, a stimulating activity may buffer against some of its demotivating influences. Regarding implications for practice, it seems beneficial to provide students with interesting and engaging tasks, which may serve as a buffer against effects of criticism. For perceived competence, the trend was reversed: the difference between negative feedback and positive was greater (larger negative effect) when tasks were deemed interesting. One explanation may be that when a feedback receiver is invested in a task, the level of perceived competence can be more adversely affected by negative feedback because performing well on an interesting task would be higher value. Students may feel worse about their abilities when being criticized, but they may still maintain interest.

### Limitations to Generalizability and Future Research

The first limitation of this synthesis is that meta-analyses in general consists of synthesis-generated evidence, which should not be interpreted as supporting causal relationships (see Cooper 1998). A synthesis can only establish an association between a moderator variable and the outcome, but not a causal connection. Second, it is important to note that the majority of these moderator tests were not significant under both fixed and random effects models, limiting the generalizability of the findings. The moderator tests were intended to be exploratory.

Third, the focus of the current synthesis was limited to feedback's effect on intrinsic motivation. While the effect of negative feedback on intrinsic motivation is important in its own right, the effect of negative feedback on performance is equally important. It has been over 20 years since the literature on the effects of negative feedback on performance was synthesized (see Kluger and Denisi 1996). As such, synthesizing the research on the feedback-performance relationship could be the focus of future syntheses of research. Taking this into account, it is important to note that intrinsic motivation can be a mediator through which feedback influences task performance. Moreover, the positive association between intrinsic motivation and performance is firmly established in the literature (Cerasoli et al. 2014; Richardson et al. 2012; Robbins et al. 2004). Thus, one can infer that if intrinsic motivation is diminished by negative feedback, then expenditure of effort may decrease followed by a decline in performance. Alternatively, if IM is bolstered by feedback, then performance could improve as a result.

Our results mirror some of the findings of Van der Kleij et al. (2015), who meta-analyzed effects of feedback on learning in computer-based environments. For instance, they found that explanatory feedback, feedback that provides an explanation, had the largest effects on performance, especially for higher order learning outcomes. In our study, we also found large effects on intrinsic motivation for feedback similar to explanatory feedback, what we call instructional feedback. Thus, although performance was not explicitly included as an outcome in our study, there are some insights to be gleaned regarding how task performance could be influenced by feedback.

In addition, a number of potentially relevant variables could not be examined as moderators of the effect of negative feedback. These omissions provide future directions for research, and we encourage researchers to explore such factors. For example, race and ethnicity have been

found in previous research to moderate the effect of criticism on motivation (Yeager et al. 2014). There is reason to believe that the effect of negative feedback may vary across cultures due to the greater value for effort over ability in more collectivist cultures (Henderlong and Lepper 2002; Shu and Lam 2016). Although we coded the sample's country origin, limited variability made it difficult to meaningfully group the samples together. The quantity and timing of feedback may also be potentially meaningful factors influencing the effects of negative feedback. However, their influence also could not be assessed as the vast majority of studies only provided feedback once, shortly after task completion. If individuals know immediately how to modify their subsequent behavior and feel autonomous and sheltered while doing so, they may have greater IM (Baron 1988). Conversely, a delay may mitigate the detrimental effect of negative feedback especially when it is threatening and controlling. Clearly, further research is needed in these areas.

One set of important variables that could not be examined in this meta-analysis were the relationship between the feedback receiver and evaluator and the qualities of the evaluator. Feedback may be received as controlling if there is mistrust or a poor relationship quality (Skipper and Douglas 2015), potentially leading to dismissal of the feedback. Without perceived sincerity of the evaluator, the extent to which the feedback can either bolster or reduce intrinsic motivation may be limited. Examining the closeness and trust between feedback receiver and giver, or even characteristics of the evaluator such as expertise, were simply not reported in the primary studies given the prevalence of experimental studies. Additionally, a growing practice in educational assessment is the use of peer feedback (i.e., Liu and Carless 2006), and further scholarship is needed to understand these varying relational dynamics in the feedback process as it relates to student motivation and learning. It is expected that the effect of negative feedback will be stronger when an evaluator with more expertise delivers it.

It is also important to note that some of the findings were based on small numbers of effect sizes, making it difficult to place a great deal of confidence in the direction and magnitude of the estimated effects. Another limitation of the study is that in all the included studies, feedback was manipulated and often bogus. That is, individuals received negative feedback regardless of how they performed on a certain task. Future research should examine the influence of authentic feedback that reflects a more accurate evaluation of the participants' performance. Similarly, due to a lack of field experiments and a majority of laboratory studies, the degree to which these effects transfer in real-world classroom settings is limited (Henderlong and Lepper 2002). For instance, unlike experiments with just one instance of feedback provided by researchers, feedback occurring in real-life academic settings is usually cumulative with multiple instances of feedback from an instructor over an academic year (Fong et al. 2018b). In contrast to the bogus feedback that was delivered to participants and related to tasks with low stakes in the reviewed literature, feedback in the real world may be also linked with high-stakes assessments. Thus, generalizability to naturalistic settings requires additional scholarship using more ecologically valid designs.

## Concluding Remarks

Under some conditions, negative feedback can reduce one's intrinsic motivation. Alternatively, with some sophistication, feedback providers can unlock student interest and persistence in a task. We have attempted to uncover the empirically documented recommendations for providing negative feedback so that it can be simultaneously motivating and instructive. Given a strong preference to solely praise when evaluating student work, teachers tend to avoid offering criticism

despite its necessary instructional implications for student learning, perhaps in fear of demotivating their students (Cohen et al. 1999). This research provides appropriate direction to guide instructors in understanding what features of negative feedback are intrinsically motivating.

Conceptually, the findings of our meta-analysis move beyond the traditional valence distinction by examining a wide range of feedback characteristics to demonstrate that negative feedback is a much more nuanced construct. Depending on the manner of delivery and other attributes of the messaging, negative feedback was shown to be not as unfavorable as the construct has been previously characterized and theorized to be like. Although our basic units of analysis were feedback statements indicating poor task performance, how the message is perceived as constructive, supportive, and hopeful can have surprising immunizing effects against negative evaluations. Perhaps negative feedback is not *negative* per se, and we encourage future work to re-consider the generalization of negatively valenced feedback. We hope this review serves to highlight and explain the rich and complex literature on negative feedback and both stimulate and direct future research addressing the role of feedback in student motivation.

## References

\*Studies included in the meta-analysis are denoted by an asterisk. Full list is in the supplement.

- Alexander, P. A., Schallert, D. L., & Hare, V. C. (1991). Coming to terms: how researchers in learning and literacy talk about knowledge. *Review of Educational Research*, 61(3), 315–343.
- Ames, C. (1992). Classrooms: goals, structures, and student motivation. *Journal of Educational Psychology*, 84(3), 261–271. <https://doi.org/10.1037/0022-0663.84.3.261>.
- \*Anderson, S., & Rodin, J. (1989). Is bad news always bad? Cue and feedback effects on intrinsic motivation. *Journal of Applied Social Psychology*, 19, 449–467. <https://doi.org/10.1111/j.1559-1816.1989.tb00067.x>, 6.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Ayduk, Ö., Gyurak, A., Akinola, M., & Mendes, W. B. (2013). Consistency over flattery: self-verification processes revealed in implicit and behavioral responses to feedback. *Social Psychological and Personality Science*, 4(5), 538–545. <https://doi.org/10.1177/1948550612471827>.
- Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. A., & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, 61(2), 213–238. <https://doi.org/10.3102/00346543061002213>.
- Baron, R. A. (1988). Negative effects of destructive criticism: impact on conflict, self-efficacy, and task performance. *Journal of Applied Psychology*, 73(2), 199–207. <https://doi.org/10.1037/0021-9010.73.2.199>.
- Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: how different are they really? *Educational Psychology Review*, 15(1), 1–40. <https://doi.org/10.1023/A:1021302408382>.
- Borenstein, M., Higgins, J. P., Hedges, L. V., & Rothstein, H. R. (2017). Basics of meta-analysis: I2 is not an absolute measure of heterogeneity. *Research Synthesis Methods*, 8(1), 5–18.
- Bracken, C. C., Jeffres, L. W., & Neuendorf, K. A. (2004). Criticism or praise? The impact of verbal versus text-only computer feedback on social presence, intrinsic motivation, and recall. *Cyber Psychology & Behavior*, 7(3), 349–357. <https://doi.org/10.1089/1094931041291358>.
- Brummelman, E., Thomaes, S., Overbeek, G., Orobio de Castro, B., Van den Hout, M. A., & Bushman, B. J. (2014). On feeding those who hunger for praise: person praise backfires among children with low self-esteem. *Journal of Experimental Psychology*, 143(1), 9–14. <https://doi.org/10.1037/a0031917>.
- Bryk, A. S., & Schneider, B. (2002). *Trust in schools: a core resource for improvement*. New York: Russell Sage Foundation.
- \*Burgers, C., Eden, A., van Engelenburg, M. D., & Buningh, S. (2015). How feedback boosts motivation and play in a brain-training game. *Computers in Human Behavior*, 48, 94–103. <https://doi.org/10.1016/j.chb.2015.01.038>.
- \*Butler, R. (1987). Task-involving and ego-involving properties of evaluation: effects of different feedback conditions on motivational perceptions, interest and performance. *Journal of Educational Psychology*, 79, 474–482. <https://doi.org/10.1037/0022-0663.79.4.474>, 4.

- \*Butler, R. (1998). Age trends in the use of social and temporal comparisons for self-evaluation: examination of a novel developmental hypothesis. *Child Development*, 69, 1054–1073. <https://doi.org/10.2307/1132362>, 4.
- Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: a theoretical synthesis. *Review of Educational Research*, 65(3), 245–281. <https://doi.org/10.2307/1170684>.
- Cerasoli, C. P., Nicklin, J. M., & Ford, M. T. (2014). Intrinsic motivation and extrinsic incentives jointly predict performance: a 40-year meta-analysis. *Psychological Bulletin*, 140(4), 980–1008.
- Cianci, A. M., Klein, H. J., & Seijts, G. H. (2010). The effect of negative feedback on tension and subsequent performance: the main and interactive effects of goal content and conscientiousness. *Journal of Applied Psychology*, 95(4), 618–630. <https://doi.org/10.1037/a0019130>.
- \*Cohen, G. L., Steele, C. M., & Ross, L. D. (1999). The mentor's dilemma: providing critical feedback across the racial divide. *Personality and Social Psychology Bulletin*, 25, 1302–1318. <https://doi.org/10.1177/0146167299258011>, 10.
- Cooper, H. (1998). *Synthesizing research: a guide for literature reviews* (3rd ed.). Thousand Oaks, CA: Russell Sage Foundation.
- Cooper, H., Hedges, L. V., & Valentine, J. C. (Eds.). (2009). *The handbook of research synthesis and meta-analysis* (2nd ed.). New York, NY: Russell Sage Foundation.
- Corpus, J. H., Ogle, C. M., & Love-Geiger, K. E. (2006). The effects of social-comparison versus mastery praise on children's intrinsic motivation. *Motivation and Emotion*, 30(4), 335–345. <https://doi.org/10.1007/s11031-006-9039-4>.
- \*Dahling, J. J., & Ruppel, C. L. (2016). Learning goal orientation buffers the effects of negative normative feedback on test self-efficacy and reattempt interest. *Learning and Individual Differences*, 50, 296–301. <https://doi.org/10.1016/j.lindif.2016.08.22>.
- De Muyneck, G., Vansteenkiste, M., Delrue, J., Altermann, N., Haerens, L., & Soenens, B. (2017). The effects of feedback valence and style on need satisfaction, self-talk, and perseverance among tennis players: an experimental study. *Journal of Sport and Exercise Psychology*, 39(1), 67–80. <https://doi.org/10.1123/jsep.2015-0326>.
- \*Deci, E. L., & Cascio, W. F. (1972). *Changes in intrinsic motivation as a function of negative feedback and threats*. Paper presented at the eastern psychological association meeting, Boston, MA.
- Deci, E. L., & Ryan, R. M. (1985a). *Intrinsic motivation and self-determination in human behavior*. New York, NY: Plenum Press.
- Deci, E. L., & Ryan, R. M. (1985b). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality*, 19(2), 109–134. [https://doi.org/10.1016/0092-6566\(85\)90023-6](https://doi.org/10.1016/0092-6566(85)90023-6).
- Deci, E. L., & Ryan, R. M. (2016). Optimizing students' motivation in the era of testing and pressure: a self-determination theory perspective. In W. C. Liu, J. C. K. Weng, & R. M. Ryan (Eds.), *Building autonomous learners* (pp. 9–29). Singapore: Springer.
- \*Deci, E. L., Cascio, W. F., & Krusell, J. (1973). *Sex differences, verbal reinforcement, and intrinsic motivation*. Paper presented at the meeting of the eastern psychological association, Washington, DC.
- Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125(6), 627–668. <https://doi.org/10.1037/0033-2909.125.6.627>.
- Duijnhouwer, H., Prins, F. J., & Stokking, K. M. (2012). Feedback providing improvement strategies and reflection on feedback use: effects on students' writing motivation, process, and performance. *Learning and Instruction*, 22(3), 171–184. <https://doi.org/10.1016/j.learninstruc.2011.10.003>.
- Duval, S., & Tweedie, R. (2000). A nonparametric “trim and fill” method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association*, 95(449), 89–98. <https://doi.org/10.1080/01621459.2000.10473905>.
- \*Dyck, D. G., Vallentyne, S., & Breen, L. J. (1979). Duration of failure, causal attributions for failure, and subsequent reactions. *Journal of Experimental Social Psychology*, 15, 122–132. [https://doi.org/10.1016/0022-1031\(79\)90024-6](https://doi.org/10.1016/0022-1031(79)90024-6), 2.
- \*Elliot, A. J., Faler, J., McGregor, H. A., Campbell, W. K., Sedikides, C., & Harackiewicz, J. M. (2000). Competence valuation as a strategic intrinsic motivation process. *Personality and Social Psychology Bulletin*, 26, 780–794. <https://doi.org/10.1177/0146167200269004>, 7.
- Fishbach, A., & Finkelstein, S. R. (2012). How feedback influence persistence, disengagement, and change in goal pursuit. In H. Aarts & A. Elliot (Eds.), *Goal-directed behavior* (pp. 203–230). New York, NY: Psychology Press.
- Fitzsimmons, G. M., & Fishbach, A. (2010). Shifting closeness: interpersonal effects of personal goal progress. *Journal of Personality and Social Psychology*, 98(4), 535–549. <https://doi.org/10.1037/a0018581>.
- Fong, C. J., Warner, J. R., Williams, K. M., Schallert, D. L., Chen, L., Williamson, Z. H., & Lin, S. (2016). Deconstructing constructive criticism: the nature of academic emotions associated with constructive,

- positive, and negative feedback. *Learning and Individual Differences*, 49, 393–399. <https://doi.org/10.1016/j.lindif.2016.05.019>.
- Fong, C. J., Schallert, D. L., Williams, K. M., Williamson, Z. H., Warner, J. R., Lin, S., & Kim, Y. W. (2018a). When feedback signals failure but offers hope for improvement: a process model of constructive criticism. *Thinking Skills and Creativity*. <https://doi.org/10.1016/j.tsc.2018.02.014>.
- Fong, C. J., Williams, K. M., Williamson, Z. H., Lin, S., Kim, Y. W., & Schallert, D. L. (2018b). “Inside out”: appraisals for achievement emotions from constructive criticism, positive feedback, and negative feedback on writing. *Motivation & Emotion*, 42(2), 236–257. <https://doi.org/10.1007/s11031-017-9658-y>.
- Gaines, L. M., Duvall, J., Webster, J. M., & Smith, R. H. (2005). Feeling good after praise for a successful performance: the importance of social comparison information. *Self and Identity*, 4(4), 373–389. <https://doi.org/10.1080/15298860500280223>.
- Greenhouse, J. B., & Iyengar, S. (1994). Sensitive analysis and diagnostics. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 383–398). New York, NY: Russell Sage Foundation.
- \*Grouzet, F. M. E., Vallerand, R. J., Thill, E. E., & Provencher, P. J. (2004). From environmental factors to outcomes: a test of an integrated motivational sequence. *Motivation and Emotion*, 28, 331–346. <https://doi.org/10.1007/s11031-004-2387-z>, 4.
- Grubbs, F. E. (1950). Sample criteria for testing outlying observations. *Journal of the American Statistical Association*, 21(1), 27–58. <https://doi.org/10.1214/aoms/117729885>.
- Harris, L. R., Brown, G. T., & Hammett, J. A. (2014). Understanding classroom feedback practices: a study of New Zealand student experiences, perceptions, and emotional responses. *Educational Assessment, Evaluation, and Accountability*, 26(2), 107–133. <https://doi.org/10.1007/s11092-013-9187-5>.
- Hattie, J. (2012). *Visible learning for teachers: maximizing impact on learning*. New York, NY: Routledge.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
- Hedges, L. V., & Vevea, J. L. (1998). Fixed and random effects models in meta-analysis. *Psychological Methods*, 3(4), 486–504. <https://doi.org/10.1037/1082-989X.3.4.486>.
- Henderlong, J., & Lepper, M. R. (2002). The effects of praise on children’s intrinsic motivation: a review and synthesis. *Psychological Bulletin*, 128(5), 774–795. <https://doi.org/10.1037/0033-2909.128.5.774>.
- Heyman, G. D., & Dweck, C. S. (1992). Achievement goals and intrinsic motivation: their relation and their role in adaptive motivation. *Motivation and Emotion*, 16(3), 231–247. <https://doi.org/10.1007/BF00991653>.
- Hirst, M. K. (1988). Intrinsic motivation as influenced by task interdependence and goal setting. *Journal of Applied Psychology*, 73(1), 96–101.
- Hu, X., Chen, Y., & Tian, B. (2016). Feeling better about self after receiving negative feedback: when the sense that ability can be improved is activated. *The Journal of Psychology*, 150(1), 72–87. <https://doi.org/10.1080/00223980.2015.1004299>.
- Ilgel, D. R., & Davis, C. A. (2000). Bearing bad news: reaction to negative performance feedback. *Applied Psychology*, 49(3), 550–565. <https://doi.org/10.1111/1464-0597.00031>.
- Ilies, R., Judge, T. A., & Wagner, D. T. (2010). The influence of cognitive and affective reactions to feedback on subsequent goals. *European Psychologist*, 15(2), 121–131. <https://doi.org/10.1027/1016-9040/a0000011>.
- \*Kamins, M. L., & Dweck, C. S. (1999). Person versus process praise and criticism: implications for contingent self-worth and coping. *Developmental Psychology*, 35, 835–847. <https://doi.org/10.1037/0012-1649.35.3.835>, 3.
- Kanouse, D. E., Gumpert, P., & Canavan-Gumpert, D. (1981). The semantics of praise. In J. H. Harvey, W. Ickes, & R. F. Kidd (Eds.), *New directions in attribution research* (pp. 97–115). Hillsdale, NJ: Erlbaum.
- Kiemer, K., Gröschner, A., Pehmer, A.-K., & Seidel, T. (2015). Effects of a classroom discourse intervention on teachers’ practice and students’ motivation to learn mathematics and science. *Learning and Instruction*, 35, 94–103. <https://doi.org/10.1016/j.learninstruc.2014.10.003>.
- Kluger, A. N., & Denisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254–284. <https://doi.org/10.1037/0033-2909.119.2.254>.
- Kulhavy, R. W. (1977). Feedback in written instruction. *Review of Educational Research*, 47(2), 211–232. <https://doi.org/10.3102/00346543047002211>.
- Kulhavy, R. W., & Stock, W. A. (1989). Feedback in written instruction: the place of response certitude. *Educational Psychology Review*, 1(4), 279–308. <https://doi.org/10.1007/BF01320096>.
- Lepper, M., & Chabay, R. W. (1985). Intrinsic motivation and instruction: conflicting view on the role of motivational processes in computer-based education. *Educational Psychologist*, 20(4), 217–230. [https://doi.org/10.1207/s15326985ep2004\\_6](https://doi.org/10.1207/s15326985ep2004_6).
- \*Lim, W. (2005). *Cognitive and motivational effects of reinforcement: a dual processing model of the impact of evaluation on creativity (unpublished doctoral dissertation)*. Indiana University, Bloomington, IN.

- Lin, L., Atkinson, R. K., Christopherson, R. M., Joseph, S. S., & Harrison, C. J. (2013). Animated agents and learning: does the type of verbal feedback they provide matter? *Computers and Education*, *67*, 239–249.
- Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., ... & Busick, M. D. (2012). *Translating the statistical representation of the effects of education interventions into more readily interpretable forms*. National Center for Special Education Research.
- Liu, N. F., & Carless, D. (2006). Peer feedback. *The learning element of peer assessment Teaching in Higher Education*, *11*(3), 279–290. <https://doi.org/10.1080/13562510600680582>.
- Locke, E. A., & Latham, G. P. (1990). Building a practically useful theory of goal setting and task motivation. *American Psychologist*, *57*(9), 707–717. <https://doi.org/10.1037/0003-066X.57.9.705>.
- Mabbe, E., Soenens, B., De Muynk, G.-J., & Vansteenkiste, M. (2018). The impact of feedback valence and communication style on intrinsic motivation in middle childhood: experimental evidence and generalization across individual differences. *Journal of Experimental Child Psychology*, *170*, 134–160. <https://doi.org/10.1016/j.jecp.2018.01.008>.
- Mangels, J. A., Rodriguez, S., Ochakovskaya, Y., & Guerra-Carrillo, B. (2017). Achievement goal task framing and fit with personal goals modulate the neurocognitive response to corrective feedback. *AERA Open*, *3*, 1–16.
- Mayer, R. E. (1996). Learners as information processors: legacies and limitations of educational psychology's second metaphor. *Educational Psychologist*, *31*(3-4), 151–161. <https://doi.org/10.1080/00461520.1996.9653263>.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & the PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(7), e1000097. <https://doi.org/10.1371/journal.pmed.1000097>.
- Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, *7*(1), 105–125. <https://doi.org/10.1037/1-82-989X.7.1.105>.
- Niemiec, C. P., & Ryan, R. M. (2009). Autonomy, competence, and relatedness in the classroom: applying self-determination theory to educational practice. *Theory and Research in Education*, *7*(2), 133–144. <https://doi.org/10.1177/1477878509104318>.
- Noddings, N. (2002). An ethic of caring and its implications for instructional arrangements. *American Journal of Education*, *96*, 215–230.
- Patall, E. A., Cooper, H., & Robinson, J. C. (2008). The effects of choice on intrinsic motivation and related outcomes: a meta-analysis of research findings. *Psychological Bulletin*, *134*(2), 270–300. <https://doi.org/10.1037/0033-2909.134.2.270>.
- Phe, G. D., & Sanders, C. E. (1994). Advice and feedback: elements of practice for problem solving. *Contemporary Educational Psychology*, *19*(3), 286–301. <https://doi.org/10.1006/ceps.1994.1022>.
- Polanin, J. R., Tanner-Smith, E. E., & Hennessy, E. A. (2016). Estimating the difference between published and unpublished effect sizes: a meta-review. *Review of Educational Research*, *86*(1), 207–236. <https://doi.org/10.3102/0034654315582067>.
- \*Pretty, G. H., & Seligman, C. (1984). Affect and the overjustification effect. *Journal of Personality and Social Psychology*, *46*(6), 1241–1253. <https://doi.org/10.1037/0022-3514.46.6.1241>.
- Raaijmakers, S. F., Baars, M., Schaap, L., Paas, F., & Van Gog, T. (2017). Effects of performance feedback valence on perceptions of invested mental effort. *Learning and Instruction*, *51*, 36–46. <https://doi.org/10.1016/j.learninstruc.2016.12.002>.
- Reeve, J. (2016). Autonomy-supportive teaching: what is it, how to do it. In W. C. Liu, J. C. K. Weng, & R. M. Ryan (Eds.), *Building autonomous learners* (pp. 129–152). Singapore: Springer.
- \*Richards, T. A. (1991). *Effects of constructive feedback on perceived competence and intrinsic motivation* (unpublished doctoral dissertation). Arizona State University, Tempe, AZ.
- Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: a systematic review and meta-analysis. *Psychological Bulletin*, *138*, 353–387. <https://doi.org/10.1037/a0026838>.
- Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, *130*, 261–288. <https://doi.org/10.1037/0033-2909.130.2.261>.
- Rosenthal, R. (1994). Parametric measures of effect sizes. In H. Cooper & L. V. Hedges (Eds.), *Handbook of research synthesis* (pp. 231–244). New York: Russell Sage Foundation.
- Rosenthal, R., & Rubin, D. B. (2003). Requaltant: a simple effect size indicator. *Psychological Methods*, *8*, 492–496.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well being. *American Psychologist*, *55*(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>.



- Ryan, R. M., Mims, V., & Koestner, R. (1983). Relation of reward contingency and interpersonal context to intrinsic motivation: a review and test using cognitive evaluation theory. *Journal of Personality and Social Psychology*, 45(4), 736–750.
- Ryan, R. M., Koestner, R., & Deci, E. L. (1991). Ego-involved persistence: when free-choice behavior is not intrinsically motivated. *Motivation and Emotion*, 15(3), 185–205.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 199–144. <https://doi.org/10.1007/BF00117714>.
- Sedikides, C., Gaertner, L., & Toguchi, Y. (2003). Pancultural self-enhancement. *Journal of Personality and Social Psychology*, 84(1), 60–79. <https://doi.org/10.1037/0022-3514.84.1.60>.
- \*Shanab, M. E., Peterson, D., Dargahi, S., & Deroian, P. (1981). The effects of positive and negative verbal feedback on the intrinsic motivation of male and female subjects. *The Journal of Social Psychology*, 115, 195–205. <https://doi.org/10.1080/00224545.1981.9711659>, 2.
- \*Shu, T., & Lam, S. (2016). Is it always good to provide positive feedback to students? The moderating effects of culture and regulatory focus. *Learning and Individual Differences*, 49, 171–177. <https://doi.org/10.1016/j.lindif.2016.06.012>.
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189. <https://doi.org/10.3102/0034654307313795>.
- Skipper, Y., & Douglas, K. (2015). The influence of teacher feedback on children's perceptions of student-teacher relationships. *British Journal of Educational Psychology*, 85(3), 276–288. <https://doi.org/10.1111/bjep.12070>.
- Steingut, R. R., Patall, E. A., & Trimble, S. S. (2017). The effect of rationale provision on motivation and performance outcomes: a meta-analysis. *Motivation Science*, 3(1), 19–50. <https://doi.org/10.1037/mot0000039>.
- Swann Jr., W. B. (2011). Self-verification theory. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (Vol. 2) (pp. 23–42). London, UK: Sage Publications.
- Swann, W. B., Pelham, B. W., & Chidester, T. R. (1988). Change through paradox: using self-verification to alter beliefs. *Journal of Personality and Social Psychology*, 54(2), 268–273.
- \*Tang, T. L. (1990). Factors affecting intrinsic motivation among university students in Taiwan. *The Journal of Social Psychology*, 130, 219–230. <https://doi.org/10.1080/00224545.1990.9924572>, 2.
- \*Tang, T. L., & Sarsfield-Baldwin, L. (1991). The effects of self-esteem, task label, and performance feedback on task liking and intrinsic motivation. *The Journal of Social Psychology*, 131, 567–572. <https://doi.org/10.1080/00224545.1991.9713886>, 4.
- Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: a social psychological perspective on mental health. *Psychological Bulletin*, 103(2), 193–210.
- Tesser, A., & Rosen, S. (1975). The reluctance to transmit bad news. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology*. New York, NY: Academic Press.
- Thorndike, E. L. (1911). *Animal intelligence: experimental studies*. Macmillan.
- \*Vallerand, R. J., & Reid, G. (1984). On the causal effects of perceived competence on intrinsic motivation: a test of cognitive evaluation theory. *Journal of Sport Psychology*, 6, 94–102, 1.
- \*Vallerand, R. J., & Reid, G. (1988). On the relative effects of positive and negative verbal feedback on males and females' intrinsic motivation. *Canadian Journal of Behavioral Sciences*, 20, 239–250. <https://doi.org/10.1037/h0079930>
- Van der Kleij, F. M., Feskens, R. C., & Eggen, T. J. (2015). Effects of feedback in a computer-based learning environment on students' learning outcomes: a meta-analysis. *Review of Educational Research*, 85(4), 475–511. <https://doi.org/10.3102/0034654314564881>.
- \*Van-Dijk, D., & Kluger, A. N. (2004). Feedback sign effect on motivation: is it moderated by regulatory focus? *Applied Psychology: An International Review*, 53, 113–135. <https://doi.org/10.1111/j.1464-0597.2004.00163.x>, 1.
- \*Viciano, J., Cervello, E. M., & Ramirez-Lechuga, J. (2007). Effect of manipulating positive and negative feedback on goal orientations, perceived motivational climate, satisfaction, task choice, perception of ability, and attitude toward physical education lessons. *Perceptual and Motor Skills*, 105, 67–82. <https://doi.org/10.2466/PMS.105.5.67-82>, 5.
- Voerman, L., Korthagen, F. A., Meijer, P. C., & Simons, R. J. (2015). Feedback revisited: adding perspectives based on positive psychology. Implications for theory and classroom practice. *Teaching and Teaching Education*, 43, 93–98. <https://doi.org/10.1016/j.tate.2014.06.005>.
- Weidinger, A. F., Steinmayr, R., & Spinath, B. (2017). Math grades and intrinsic motivation in elementary school: a longitudinal investigation of their association. *British Journal of Educational Psychology*, 87(2), 187–204.
- Weiner, B. (2014). The attribution approach to emotion and motivation: history, hypotheses, home runs, headaches/heartaches. *Emotion Review*, 6(4), 353–361. <https://doi.org/10.1177/1754073914534502>.

- Wicker, F. W., Brown, G., & Paredes, V. (1990). Competing activities and measures of intrinsic motivation. *Journal of Social Psychology, 130*(6), 813–819. <https://doi.org/10.1080/00224545.1990.9924633>.
- Williams, K. J., Donovan, J. J., & Dodge, T. L. (2010). Self-regulation of performance: goal establishment and goal revision processes in athletes. *Human Performance, 13*(2), 159–180. [https://doi.org/10.1207/s15327043hup1302\\_3](https://doi.org/10.1207/s15327043hup1302_3).
- Winstone, N. E., Nash, R. A., Parker, M., & Rowntree, J. (2017). Supporting learners' agentic engagement with feedback: a systematic review and a taxonomy of recipience processes. *Educational Psychologist, 52*(1), 17–37. <https://doi.org/10.1080/00461520.2016.1207538>.
- \*Woodcock, A. J. (1990). *The effect of verbal feedback on intrinsic motivation and perceived competence of cricketers* (unpublished doctoral dissertation). Arizona State University, Tempe, AZ.
- \*Yeager, D. S., Purdie-Vaughns, V., Garcia, J., Apfel, N., Brzustoski, P., Master, A., Hessert, W. T., Williams, M. E., Cohen, G. L. (2014). Breaking the cycle of mistrust: wise interventions to provide critical feedback across the racial divide. *Journal of Experimental Psychology, 143*, 804–824. <https://doi.org/10.1037/a0033906>, 2.