Beyond cold technology: a systematic review and meta-analysis on emotions in technology-based learning environments

Kristina Loderer, Reinhard Pekrun, James C. Lester

Angaben zur Veröffentlichung / Publication details:

Loderer, Kristina, Reinhard Pekrun, and James C. Lester. 2020. "Beyond cold technology: a systematic review and meta-analysis on emotions in technology-based learning environments." *Learning and Instruction* 70: 101162.
https://doi.org/10.1016/j.learninstruc.2018.08.002.

Beyond cold technology: A systematic review and meta-analysis on emotions in technology-based learning environments[☆]

Kristina Loderer[a,*], Reinhard Pekrun[a,b], James C. Lester[c]

[a] Department of Psychology, University of Munich, Munich, Germany
[b] Institute for Positive Psychology, Australian Catholic University, Sydney, NSW, Australia
[c] Department of Computer Science, North Carolina State University, Raleigh, NC, USA

1. Introduction

Understanding and supporting emotional processes involved in technology-based learning (TBL) has become a paramount goal across different research communities focused on different types of environments (TBLEs) such as online content-management platforms (e.g., Artino, 2009), hypermedia systems (e.g., Cromley, Azevedo, & Olson, 2005), virtual realities (e.g., Noteborn, Bohle Carbonell, Dailey-Hebert, & Gijselaers, 2012), or intelligent tutoring systems (ITS; e.g., Graesser, Chipman, King, McDaniel, & D'Mello, 2007). From a broader perspective, the roots of this burgeoning field of research can be traced back to the educational technology revolution in the 1970s. During this early stage, causes and effects of anxiety in drill-and-practice computer-assisted instruction were studied extensively (Sieber, O'Neil, & Tobias, 1977). Subsequently, scholars turned to correlates of "technophobia" in the context of educational-vocational technology training in parallel to the computer literacy movement in the 1980s (Moreno, 2012).

This early focus on anxiety in TBL settings is not too surprising when considering that, educational emotion research was almost exclusively focused on test anxiety for nearly half a century since its beginnings in the 1930s. This intense devotion resulted in the development of a number of influential theories explaining the antecedents and consequents of this emotion. Early conceptual models include psychoanalytic approaches to achievement-related anxiety (e.g., Stengel, 1936), neo-behavioristic theories (e.g., Mandler & Sarason, 1952; McKeachie, 1951), and motivationally-focused approaches to fear of failure (e.g., Atkinson & Feather, 1966; see also Hagtvet & Benson, 1997). From the early 1970s onwards, a multitude of cognitive theories seeking to explain the test anxiety phenomenon was developed, including, for instance, Wine's cognitive-attentional model (1971), appraisal-based transactional models of stress and emotions (Lazarus & Folkman, 1984; see also; Spielberger & Vagg, 1995), Covington's (1984) self-worth model, and socio-cognitive expectancy-value approaches to (e.g., Pekrun, 1992). Building on and expanding on these concepts, recent theoretical advances in modeling origins and consequents of test anxiety include person-situation interactional approaches such as the self-referent executive function model of emotional distress (Zeidner & Matthews, 2005; see Putwain, 2008, as well as Zeidner, 1998, 2014, for detailed reviews).

However, as psychological research on the role of emotions (beyond anxiety) in human cognition and motivation increasingly caught the interest of educational scientists during the 1990s (Pekrun, 2005), TBL

☆ This research was supported by a grant from the Learning Environments Across Disciplines (LEADS) network (Canadian Social Sciences and Humanities Research Council, [SSHRC] partnership grant 895-2011-1006, 2011) awarded to the second author.

* Corresponding author. Department of Psychology, University of Munich, Leopoldstrasse 13, D-80802, Munich, Germany.
E-mail address: Kristina.Loderer@psy.lmu.de (K. Loderer).

researchers too turned to a broader array of emotions and their linkages with learning. This expansion was guided by a range of theories, including theories of self-regulated learning (e.g., Azevedo, Johnson, Chauncey, & Burkett, 2010), interest theory (Hidi & Renninger, 2006), approaches considering the influence of impasses and cognitive incongruity on emotions (D'Mello & Graesser, 2014; Graesser, D'Mello, & Strain, 2014), models of technology acceptance (Davis, 1989), the cognitive-affective theory of learning with media (Moreno, 2006), Plass and Kaplan's (2016) integrated cognitive affective model of multimedia learning, and the control-value theory of achievement emotions (CVT; Pekrun, 2006; Pekrun & Perry, 2014). As a result, researchers using these theories have operated in relative isolation, despite their common interest in the role of emotions in TBL.

The CVT offers an exemplar for an integrative approach to emotions and learning. It integrates propositions from appraisal theories (Shuman & Scherer, 2014), expectancy-value models of emotions (Turner & Schallert, 2001), transactional theories of stress-related emotions (Lazarus & Folkman, 1984), attribution theory (Graham & Taylor, 2014; Weiner, 2007), and models addressing the effects of emotions on learning and performance (Fredrickson, 2001; Zeidner, 1998) to provide a platform for research on emotions and learning across different research paradigms and educational environments. Employing this theory as a framework, the present review synthesizes empirical findings of nearly five decades of research on antecedents and outcomes of emotions in TBLEs in the form of a two-stage systematic review and selective meta-analysis (see section 1.2 for details). By integrating extant research, we seek to examine whether emotional mechanisms of learning generalize across learner populations, subject domains, research methodologies, and different types of TBLEs to build a basis for more integrative perspectives. Our approach is grounded in the relative universality proposition of the CVT which predicts differences in mean levels of emotions between learning environments, but generality of functional mechanisms pertaining to antecedents and effects of emotions. This approach also allows for examining whether emotional mechanisms of TBL are consistent with those reported for more traditional, non-technology-based forms of learning, catering to calls for more unified accounts of emotional facets of learning and the deduction of evidence-based principles of affective design of TBLEs.

1.1. The role of emotions in TBL

In a meta-analysis on incidence rates of emotions across different TBLEs, D'Mello (2013) illustrated that learners' affective experiences in these environments are, like learning more generally, highly multifaceted, covering a range of positive and negative emotions such as enjoyment, curiosity, anxiety, anger, confusion, and boredom. Drawing on a host of established theoretical accounts of emotions, the CVT provides one helpful framework for systematizing evidence for antecedents and outcomes of these emotions. Its core propositions targeting achievement emotions tied to learning activities and success/failure outcomes, and recent extensions to other groups of emotions relevant to TBL, are reviewed below and detailed in Table 1.

1.1.1. Appraisal antecedents of emotions during TBL

Perceived control and value of achievement activities and outcomes form important antecedents of achievement emotions (for summaries of supporting evidence, see Graham & Taylor, 2014; Hembree, 1988; Pekrun & Perry, 2014; Zeidner, 1998, 2014). Achievement emotions are thought to be instigated when the individual feels in control over, or out of control over, subjectively important achievement activities or outcomes. Perceived control pertains to one's perceived ability to effectively manage a given achievement situation, as implied by causal expectations and attributions of success and failure (e.g., Weiner, 2007) as well as underlying competence beliefs (e.g., self-concept of ability; Marsh, 1993). Perceived value includes both valence (positive vs. negative) and subjective importance (e.g., intrinsic interest or instrumental usefulness) of achievement activities and outcomes (see Wigfield, Rosenzweig, & Eccles, 2017, for a recent review). It is hypothesized that perceived control positively influences positive emotions, and negatively influences negative emotions (e.g., King & Gaerlan, 2014; Niculescu, Tempelaar, Dailey-Hebert, Segers, & Gijselaers, 2015; Putwain, Sander, & Larkin, 2013), except for boredom. Boredom is experienced when perceived control is either too low (over-challenge) or too high (under-challenge; e.g., Acee et al., 2010). Value is held to amplify both positive and negative emotions. Specifically, positive value of learning and achieving is thought to trigger positive and reduce negative emotions and negative value focused on failure or inability to master a task is held to increase the intensity and frequency of negative emotions like anxiety or anger (Bieg, Goetz, & Hubbard, 2013; Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011).

The CVT proposes that functional mechanisms of human emotions are bound to universal, species-specific characteristics of our mind (Pekrun, 2006). As such, it suggests that these appraisal patterns should be stable across individuals, genders, academic domains, socio-cultural contexts, and also different learning environments, including traditional learning environments as well as different TBLEs. Supporting this assumption, Daniels and Stupnisky (2012) contend that “although there are different targets for the appraisals of control and value when course delivery changes, ultimately, students are still evaluating their levels of perceived control and task value … in the end both appraisals appear to affect the experience of discrete achievement emotions in much the same way as has been found in face-to-face classrooms” (p. 225). Similarly, recent research shows that control-value perceptions are also relevant to the arousal of epistemic emotions tied to knowledge-generating qualities of cognitive tasks (e.g., surprise, curiosity, and confusion; Muis, Chevrier, & Singh, 2018; Pekrun, Vogl, Muis, & Sinatra, 2017) as well as emotions targeted directly at the technology used (Butz, Stupnisky, & Pekrun, 2015). For these epistemic and technology emotions, control-value appraisal constellations should follow similar patterns as those of achievement emotions. Specifically, positive emotions like epistemic curiosity or technology-related enjoyment should be linked to perceptions of high control and positive valuation of the cognitive task or technology at hand, respectively. Negative emotions such as frustration or boredom in response to cognitive tasks or the technology itself should relate negatively to learners' perceived control and positive value (Butz et al., 2015; Muis, Psaradellis, Lajoie, Di Leo, & Chevrier, 2015; see also; Loderer, Pekrun, & Plass, in press). Table 1 summarizes hypothesized relations between control-value appraisals and emotions considered in this synthesis.

1.1.2. Characteristics of learners and environments as distal antecedents

Differences between learners and learning environments are posited to shape learners' emotions by impacting their perceptions of control and value. For instance, learners' prior knowledge in a domain (e.g., math) or experience with technology and TBLEs more generally may influence their perceptions of control over a given learning task, thus impacting their emotions. Differences may also be linked to students' gender. Prevailing stereotypes about domain-specific gender differences in math and language abilities as well as technology use can impact domain-related competence beliefs, thus contributing to gender differences in emotions (e.g., math emotions; Beilock, Gunderson, Ramirez, & Levine, 2010; Goetz, Bieg, Lüdtke, Pekrun, & Hall, 2013). Meta-analytic evidence also suggests that females are prone to higher levels of test anxiety (Hembree, 1988; Seipp & Schwarzer, 1996). However, evidence indicates that computer literacy gaps between males and females are closing (Gunn, French, McLeod, McSporran, & Conole, 2002), and that stereotypically male videogame genres including competitive fighting games have gained favor with girls as well. This suggests that traditional gender differences in technology use may be receding (Homer, Hayward, Frye, & Plass, 2012).

Similarly, control-value appraisals are held to mediate the impact of the learning environment on learners' emotions. Two important factors

Table 1
Summary of correlates and hypothesized relations with emotions.

Construct	Common indicators	Emotion			Potential moderator effects
		Positive Activating	*Negative Activating*	*Negative Deactivating*	
		enjoyment, curiosity/ interest	anxiety, anger/ frustration, confusion	boredom	
Perceived control	self-concept; self-efficacy; perceived ease of use	+	–	+/− (likely -)	none
Perceived value	intrinsic or utility value; perceived usefulness; task value	+ for pos. value - for neg. value	- for pos. value + for neg. value	- for pos. value + for neg. value	none
Gender	–	±	±	±	type of TBLE[a]; emotion focus; culture
Prior knowledge	pre-test score; prior achievement; number of prior courses	+	–	±	none
(Meta-)cognitive support	provision of hints, guidance or learning aids (e.g., visual organizers; agent messages)	+	–	–	type of TBLE[a]
Cognitive conflict	contradictory information; false feedback	±	+	±	type of TBLE[a]
Peer/teacher esteem	perceived or manipulated other-expressed importance/appreciation of a task/technology	+ for pos. value - for neg. value	- for pos. value + for neg. value	- for pos. value + for neg. value	type of TBLE[a]
Aesthetic design	increasing aesthetic appeal through color, sound, visual illustrations, or animation	+	–	–	type of TBLE[a]
TBLE quality	subjective quality of TBLE (e.g., overall satisfaction)	+	–	–	none
TBLE vs. non-TBLE	comparison of TBLE[a] to an active control group learning the same content without technology	±	±	±	type of TBLE[a]; emotion focus; gender; culture
Engagement	intrinsic, extrinsic, or general academic motivation; behavioral engagement; flow	+	+/− (likely -)	–	none
Disengagement	off-task behavior; gaming the system; procrastination	–	+/− (likely +)	+/− (likely +)	none
Strategy use	elaboration; metacognitive strategies; rehearsal; combinations of strategies	+	+/− (likely -)	–	none
Learning outcomes	task-/course-level grades; number of errors (reverse-coded)	+	+/− (likely -)	–	none

Note. '+' indicates positive and '-'negative relations; '+/−' indicates that variable effects are possible. The categories 'engagement' and 'disengagement' each included both motivational (e.g., intrinsic motivation or amotivation to learn, respectively) and behavioral indicators (e.g., effort investment or off-task behavior, respectively).

[a] Moderating effects are possible to the extent that environments differ in their effectiveness. Due to data constraints, this was not tested (see discussion/ conclusions).

also considered in McNamara, Jackson, and Graesser's (2010) intelligent tutoring and games framework (ITaG) are the cognitive and motivational qualities of TBLEs (see also Plass & Kaplan, 2016). Cognitive quality comprises clarity and structure of tasks which may influence learners' perceptions of control over, and value of a given task or technological tool. Providing cognitive and metacognitive support to promote clarity may thus be an effective way to foster positive and reduce negative emotions (Goetz et al., 2013; Plass & Kaplan, 2016). Inducing cognitive conflicts through contradictory information, in turn, has been shown to evoke confusion (D'Mello & Graesser, 2014). Motivational qualities that communicate value of learning technology are also considered important for students' emotions. Such values may be communicated by students' peers or instructors, for instance, through messages that convey information on the utility of learning contents or technology tools (Frenzel, Pekrun, & Goetz, 2007b; see also Gaspard et al., 2015, on the related concept of social utility value). Factors that enhance the aesthetic appeal of learning through visual illustrations or auditory enhancements are expected to the increase intrinsic value of learning material and thereby foster positive emotions such as curiosity and enjoyment (Loderer et al., in press; Plass & Kaplan, 2016).

From an emotional perspective, these cognitive and motivational characteristics are relevant to all learning environments. Their realizations, however, may vary with the type of environment considered. For example, metacognitive support can be supplied by agents or through visualized cognitive maps that provide an outline of learning contents (Beasley & Waugh, 1995). From the perspective that in both cases, the different environmental features target common mechanisms in terms of learners' control-value appraisals, they should yield similar emotional consequences, but variability may be expected to the degree that interventions differ in their effectiveness to produce the desired cognitive and motivational qualities.

In terms of overall learning modality, the question of technology-based and non-technology-based variants of learning should differ in mean levels of positive and negative learner emotions is difficult to answer. It is tempting to assume that students will automatically be more engaged with visually stimulating 3D animations (but see Moos & Marroquin, 2010, for a critical perspective). From the perspective that different educational environments may fulfill learners' needs to similar extent, albeit through different means, however, a simple 'yes' implying that TBLEs are inherently more enjoyable seems unjustified, implying that modality comparisons may have variable effects on emotions (Table 1).

1.1.3. Functions of emotions for learning and achievement in TBLEs

Emotions are thought to influence important components of learning processes, such as attention, motivation, and use of learning strategies, as well as resulting learning outcomes (Fredrickson, 2001; Pekrun & Perry, 2014; Zeidner, 1998). Achievement-related, epistemic, and technology emotions likely influence these processes in similar ways. For example, frustration with a seemingly unsolvable learning task likely bears similar motivational consequences for achievement as frustration targeted at the technological device used for learning. In their review of classroom research on emotions, Goetz and Hall (2013) observed that relations between discrete learning-related emotions (enjoyment, pride, anxiety, anger, boredom) and achievement across primary, secondary, and tertiary education average at approximately

$r = |0.25|$ (range = $|0.04$ to $0.40|$). Positive activating emotions such as curiosity and enjoyment (Table 1) are posited to promote both intrinsic and extrinsic motivation as prerequisite for effort investment, as well as strategic forms of learning and use of flexible, deep learning strategies, thus positively influencing learning outcomes under most conditions (e.g., Ahmed, van der Werf, Kuyper, & Minnaert, 2013; Ranellucci, Hall, & Goetz, 2015). Negative deactivating emotions such as boredom (Table 1), in contrast, are expected to reduce motivation and effortful, strategic learning, thereby negatively impacting achievement (Pekrun, Hall, Goetz, & Perry, 2014; Tze, Daniels, & Klassen, 2016). Research suggests that boredom and other negative emotions may be linked to motivational disengagement in TBLEs, as indicated by off-task behavior or gaming the system (i.e., completing tasks by taking advantage of system properties; Sabourin & Lester, 2014).

Effects of negative activating emotions (e.g., anxiety, anger, frustration, confusion; Table 1) on learning and performance are thought to be more variable, given that they undermine intrinsic motivation to learn or interact with a TBLE but can also induce extrinsic motivation to avoid failure and foster more rigid, rehearsal-based learning. To the extent that algorithmic thinking or memorization is conducive to specific task requirements, effects on achievement may be positive for these emotions (see also Goetz & Hall, 2013).

Recent work on epistemic emotions also suggests that confusion may promote complex thinking and foster learning if resolvable (D'Mello & Graesser, 2014). However, especially for emotions like anxiety or frustration, negative effects on overall learning outcomes likely outweigh any beneficial short-term effects for most learners (Pekrun et al., 2011). In line with this reasoning, meta-analytic evidence documents moderate negative relations ranging from $r = -0.20$ to -0.25 between anxiety and various indicators of academic performance (Hembree, 1988; Seipp, 1991; see also Zeidner, 2014, for a recent review).

1.2. Aims of the present review

As outlined above, research on emotions in TBLEs has been guided by a range of theoretical approaches. As Plass and Kaplan (2016) illustrate, however, their propositions regarding antecedents and effects of emotions during learning are often aligned or complementary. Furthermore, in their review of a special issue on emotions in online courses, learning with virtual realities, or intelligent tutoring systems, Daniels and Stupnisky (2012) proposed that emotional mechanisms of TBL and non-TBL are 'not that different in theory', implying that the CVT can serve as a platform for studying emotions and learning across research disciplines and educational settings.

This review aims to expand these initial synergistic efforts through a CVT-based integration of evidence for antecedents and effects of emotions in TBLEs. Our analysis of extant findings comprised two stages. During the first stage which was completed in 2014, we systematically surveyed research published via various publication outlets until May 2014 to provide an overview over the types of antecedents and effects of emotions that have been investigated in TBL settings. Furthermore, we statistically integrated empirical findings for emotion-correlate relations examined by sufficient a number of studies warranting meta-analytic synthesis. During the second stage completed in 2018, we searched for and coded empirical work published between May 2014 and July 2018 addressing these same emotion-correlate relations. We compare these additional findings with the effect sizes collected during the first stage to probe the consistency of findings published prior to and after May 2014. Furthermore, we aim to identify for which relations the number of available studies was particularly limited and use this information to highlight directions for future research in this emergent field.

Our approach is grounded in the relative universality proposition of the CVT which predicts differences in mean levels of emotions between learners, genders, cultures, or learning environments, but generality of functional mechanisms pertaining to antecedents and effects of emotions across these variables (Pekrun, 2009). Specifically, we synthesized research on relations between emotions (enjoyment, curiosity/interest, anxiety, anger/frustration, confusion, boredom) and their antecedents (control-value appraisals, prior knowledge, gender) as well as outcomes (motivational engagement, use of learning strategy, learning outcomes) within TBLEs. We expected correlations to parallel those reported in classroom-based research (Table 1), especially with regard to the direction of effects. Furthermore, we examined whether different characteristics of TBLEs (provision of [meta-] cognitive support, induction of cognitive conflict, aesthetic design, peer/instructor esteem of learning, perceived quality of TBLE) produce expected mean level differences in positive and negative emotions. We also compiled direct comparisons of levels of emotions in TBL versus non-TBL. Based on the previous deliberations and the relative universality proposition, gender composition, cultural background, type of TBLE, and object focus of emotion were examined as focal moderators of meta-analytically computed mean effects obtained during the first stage of the review (see section 3.4). To the extent that relative universality holds true, effect sizes for linkages between emotions and appraisal and prior knowledge antecedents as well as outcomes should remain consistent across different levels of these moderators.

2. Method

The present review comprised two stages of study searching, screening, and coding. The initial review stage (stage 1) covered research on emotions in TBLEs published until May 2014. Based on a systematic screening of this research, we identified and analyzed the most commonly investigated emotion-correlate relations across the empirical (i.e., quantitative) studies available. In a second step (stage 2), we specifically searched for and extracted effect sizes quantifying the selected relations from empirical studies published after May 2014 to examine the robustness of the findings over time. Study selection and coding for both review stages are detailed in sections 2.1 to 2.3.

2.1. Inclusion criteria

2.1.1. Initial review (stage 1)

Studies had to meet seven criteria to be included. First, they include directly measured or experimentally induced discrete emotions relevant to learning (Pekrun & Stephens, 2012). To ensure consistency in eligibility decisions and coding, working definitions of emotion categories initially considered for inclusion were derived (see Supplementary Material S1). In line with current interest theories (Ainley & Hidi, 2014; Silvia, 2008), interest was included in the 'curiosity/interest' category only if it was operationalized as a state-level experience within a single learning session (i.e., situational interest). Anger and frustration were also combined into one category as primary studies often combined both within single scales. Studies had to provide sufficient information to judge alignment with our definitions in the form of sample items, observation protocols, or references to accessible sources, or have relied on previously validated instruments and induction procedures. Single item measures were to explicitly label the target emotion. Bipolar items mixing discrete categories (e.g., boring – exciting) were excluded. Similarly, we excluded studies in which emotions were solely inferred from physiological data, EEG, body posture, gaze, or discourse due to open issues pertaining to the validity of these measures as indicators of emotion (D'Mello, 2013).

Second, studies had to involve hands-on interaction with technology for an educational purpose on the part of the learner in either field or lab settings. Thus, we included studies that investigated (1) learners' acquisition of technology skills (e.g., programming) or learning with (2) web/computer tools (e.g., text annotation software), (3) learning/content management systems (LMS/CMS) that provide platforms for instruction, communication, and assessment in in online/hybrid courses, (4) non-intelligent learning programs focused on content delivery through multiple forms of representation, (5) intelligent tutoring systems involving computational modeling of and adaptation to learner variables (e.g., learning progress), and/or (6) virtual/augmented realities providing lifelike simulations of real or imaginary worlds (Lajoie & Azevedo, 2006; Moos & Azevedo, 2009). Studies in which teachers were learning to use technology were included, while those in which technology merely supplemented teacher-based instruction or studies on computer-based testing only were excluded.

Third, studies had to have measured or experimentally manipulated at least one emotion-antecedent or -consequent variable relevant to the CVT. The list of initially considered correlates was narrowed down systematically during study screening (see 1.2). Fourth, studies had to be based on nonclinical samples (i.e., excluding studies on computerized therapy of affective disorders or learning), and fifth, be published in English in the form of a peer-reviewed article. Sixth, studies had to provide effect size information or original data for calculating an effect size for at least one emotion-correlate relation of interest. Seventh, effect sizes were coded only if correlates pertaining to antecedents of emotions (i.e., control-value appraisals, prior knowledge/experience, TBLE characteristics) were measured prior to, or at approximately the same time as the emotion, and if correlates pertaining to learning processes and outcomes (i.e., motivational variables, learning strategies, achievement) were measured approximately at the same time or after the emotion.

2.1.2. *Supplemental review (stage 2)*

Study inclusion was determined using the same criteria as used for the initial review. However, as noted, only studies providing effect sizes or adequate data for the emotion-correlate relations identified in the initial review were retained for coding.

2.2. *Literature search, eligibility screening, and coding*

2.2.1. *Initial review (stage 1)*

The databases PsycINFO and ERIC were searched using a strategy combining multiple terms related to emotions, learning/instruction, and technologies (see Table S2 for a complete list). This search was concluded on May 31, 2014 and yielded 1040 unique records for the years from 1952 to 2014. Abstracts were pre-screened for basic mention of criteria (i.e., emotions, technology, learning, quantitative methodology) by two raters with 97% agreement on preliminary eligibility. At this stage, 365 studies were retained. Second, we additionally searched (a) peer-reviewed proceedings of the Intelligent Tutoring Systems, Artificial Intelligence in Education, and Educational Data Mining conferences, (b) specific peer-reviewed journals (*International Journal of Learning Technology, IEEE Transactions on Affective Computing, Journal of Educational Data Mining*), and (c) reference lists of an edited book (Calvo & D'Mello, 2011) and a chapter (Graesser, D'Mello, & Strain, 2014) for potential studies published up to May 2014. An additional 276 relevant records were retrieved.

Full texts of the remaining 641 records were independently screened by two raters and simultaneously coded for emotion constructs. Agreement on study inclusion was 87%, and disagreement settled through discussion with the second author. A total of 172 studies were retained for preliminary coding. Agreement on emotion constructs accepted by both raters (n = 149) ranged from 89% to 93% across emotion categories. After reaching agreement on the remaining studies, anxiety, enjoyment, anger/frustration, boredom, confusion, and curiosity/interest were the only emotions identified as measured by at least 10 independent studies (106, 65, 25, 23, 21, and 15 studies, respectively) and selected for the analysis.

Next, two raters collaboratively scanned these 172 studies for investigated emotion-correlate relations. Finally, correlates measured in connection to at least one of the six target emotions by at least five independent studies were retained (see 2.4; Table 1), yielding a final sample of 149 studies for inclusion.

The final coding scheme included descriptors of the emotions, correlates, and TBLEs examined by each study, as well as effect size information (see Supplementary Material for additional information). Specifically, we included items pertaining to how emotions and correlates were measured (state-vs. trait-level; self-report vs. behavioral measures or combinations of both) or induced. Where given, we coded reliabilities (internal consistency, kappas). Emotions were classified as (1) technology-focused, (2) relating to learning (contents) or achievement (altogether labelled ‘academic), or (3) unspecified/mixed (e.g., simply asking learners to report current feelings). We also coded type of TBLE (see 2.1). Finally, we added several sample descriptors (age group, gender composition), subject domain, and cultural background (continent). Two raters independently coded 40 randomly selected studies, also noting relevant data for effect size coding. After resolving coding discrepancies, rater 1 (first author) coded all 149 studies. Study-level coding details are provided in Table S3.

2.2.2. *Supplemental review (stage 2)*

Because database searching constituted the most prolific source for locating eligible studies for our initial review (121 of 149 studies; 81.2%), we conducted an additional database search for peer-reviewed studies indexed in PsycINFO or ERIC and published between June 2014 and July 2018. The search strategy included the same terms employed for the initial literature search (stage 1), the only exception being that the emotion terms were restricted to the six emotions selected for inclusion in our initial review (i.e., anxiety, enjoyment, anger/frustration, boredom, confusion, and curiosity/interest; see Table S2 for details). This search yielded 461 unique records. Abstracts were pre-screened based on the same criteria as applied during the initial review (see 2.1.1) by the first author. Full-texts of 179 accessible articles were screened in detail for inclusion in the supplemental review, yielding a final sample of 37 additional studies for inclusion. Studies were coded by the first author using the coding scheme developed during the initial review stage (see 2.1.1). Study-level coding details are provided in Table S8.

2.3. *Effect size calculation*

For all studies included in both stages of the review, emotion-correlate relations were expressed in terms of correlation coefficients as these account for the continuous nature of the majority of variables and relations examined (e.g., degree of perceived task value and intensity of emotional experience). Statistically, correlations can be readily used to represent both correlational and (quasi-)experimental designs in which groups are contrasted on continuous (e.g., intensity of emotional experience) or categorical outcomes (e.g., high versus low emotional experience), allowing for meaningful comparison across studies (Schmidt & Hunter, 2015). Effect sizes and sampling variances were directly extracted from primary studies or calculated from descriptive and inferential statistics using the equations in Table S3. Care was taken to ensure comparability of metrics across study designs.

Where studies yielded multiple rs for a given relation, for instance, due to having administered multiple achievement tests, all were extracted. Data from multivariate analyses (e.g., multiple regression, structural equation modeling) or partial correlations were excluded due to lack of comparability across studies (Lipsey & Wilson, 2001). To counteract biasing results by excluding studies reporting that effects were significant or nonsignificant without reporting effect sizes, we randomly generated plausible effect sizes within the range of the reported information (Murayama, Miyatsu, Buchli, & Storm, 2014; Table S3). For relations between emotions and learning outcomes, and studies comparing emotions in TBLEs versus non-TBLEs, several studies allowed for computing adjusted effect sizes (i.e., adjusted for prior knowledge, or baseline differences in emotions, respectively). In these cases, both adjusted and unadjusted effect sizes were extracted but analyzed separately. For additional details on effect size computations, see Supplementary Material (Method) and Tables S3,S4, and S8.

2.4. Statistical analyses

As part of our initial review (stage 1), we sought to integrate empirical findings for emotion-correlate relations examined by a statistically sufficient number of primary studies. As such, the following sections 3.4.1 to 3.4.3 delineate analyses computed based on the 149 studies included in the initial review. Findings collected in the supplemental review were examined descriptively (section 4.3).

2.4.1. Estimating mean and moderator effects

As the 149 primary studies included in our initial review differed in methodological characteristics that likely introduce between-study variation in effect sizes, a random-effects framework was adopted (Hedges & Vevea, 1998). All analyses were performed in R (version

Table 2
Mean weighted effect sizes ($\bar{r}$) for relations between emotions, antecedents, and outcomes computed in the initial review (stage 1).

Correlate	Emotion	k (n_{ES})	n_s	$\bar{r}$	95% CI		Range (r)	τ^2
					LL	UL		
Antecedents								
Appraisal								
Perceived control	enjoyment	12 (15)	2188	.50***	.38	.60	.22–.76	.04
	anxiety	25 (43)	4792	-.27***	-.38	-.16	-.86–.52	.05
	outliers removed	24 (40)	4772	-.26***	-.36	-.16	-.68–.07	.04
Perceived value	enjoyment	14 (19)	2802	.56***	.44	.66	.22–.84	.07
	anxiety	17 (28)	2449	-.13*	-.24	-.02	-.54–.23	.04
Learner characteristics								
Gender	enjoyment	6 (8)	1188	-.04[†]	-.15	.07	-.18–.29	.00
	outliers removed	5 (7)	1158	-.05[†]	-.18	.08	-.18–.04	.00
	anxiety	21 (28)	2535	-.04	-.03	.12	-.38–.31	.02
Prior knowledge	enjoyment	8 (18)	776	.20*	.06	.33	-.06–.46	.01
	anxiety	24 (68)	2448	-.28***	-.36	-.21	-.66–.24	.03
TBLE characteristics								
(Meta-)cognitive support	enjoyment	9 (10)	1114	.20**	.09	.31	-.02–.44	.00
	confusion	6 (6)	299	-.19[†]	-.47	.13	-.59–.12	.04
Cognitive conflict	confusion	5 (9)	379	.07[†]	-.01	.14	-.10–.20	.01
Peer/instructor esteem	anxiety	6 (6)	808	-.16*	-.31	-.01	-.36–.04	.01
Aesthetic design	curiosity/interest	8 (10)	416	.17**	.08	.26	-.02–.30	.00
Perceived TBLE quality	enjoyment	8 (12)	840	.52***	.37	.65	-.02–.79	.04
	anxiety	7 (16)	763	-.28*	-.49	-.03	-.76–.45	.05
	outliers removed	5 (12)	697	-.17[†]	-.31	-.03	-.24–.01	.00
TBLE vs. non-TBLE comparison[a]	enjoyment	6 (9)	268	.16	-.19	.48	-.40–.51	.10
	anxiety	12 (20)	1488	-.07	-.16	.02	-.29–.19	.02
	adjusted	8 (16)	781	-.13	-.32	.07	-.49–.32	.07
Outcomes								
Learning processes								
Engagement	enjoyment	9 (14)	1872	.45**	.20	.64	.06–.80	.13
	anxiety	14 (31)	2962	-.12*	-.23	-.01	-.55–.17	.03
Disengagement	enjoyment	5 (8)	546	-.10[†]	-.17	-.03	-.20–.14	.00
Strategy use	enjoyment	7 (18)	2022	.31**	.13	.48	-.15–.53	.05
	anxiety	15 (30)	2464	-.14*	-.24	-.02	-.66–.41	.04
	outliers removed	15 (28)	2268	-.17**	-.26	-.05	-.58–.20	.03
Learning outcomes	enjoyment	11 (20)	1458	.15	-.00	.30	-.32–.73	.04
	outliers removed	10 (18)	1384	.18**	.10	.26	-.17–.35	.01
	curiosity/interest	5 (8)	638	.20[†]	.07	.32	.13–.35	.00
	anxiety	30 (51)	4049	-.17***	-.24	-.09	- 1–.22	.03
	outliers removed	29 (48)	4031	-.14***	-.21	-.08	-.54–.22	.02
	anger/frustration	6 (8)	312	-.07	-.32	.19	-.46–.15	.03
	confusion	7 (11)	400	-.09	-.44	.28	-.67–.49	.10
	boredom	8 (14)	1181	-.08	-.21	.05	-.41–.24	.02

Note. For relations between emotions and appraisals, learner characteristics, outcome variables, and TBLE quality, positive *r*s reflect higher levels of emotions being connected to higher levels of the respective correlate, and negative *r*s reflect higher levels of emotions connected to lower levels of the correlate. For gender, positive *r*s indicate higher levels of the respective emotion for females compared with males. For TBLE characteristics, positive *r*s reflect (meta-)cognitive support, cognitive conflict, peer/teacher esteem of learning, aesthetic design, and TBLE versus non-TBLE being associated with higher levels of the respective emotion examined, while negative *r*s indicate lower levels of the respective emotions. k = number of independent studies, n_{ES} = number of effect sizes. LL = lower limit, UL = upper limit of 95% confidence interval (CI). τ^2 = between-study heterogeneity of effect sizes.
***$p < .001$. **$p < .01$. *$p < .05$..[(†)]p-value not reliable ($df < 4$ for t-test of $\bar{r}$).
[a] $\bar{r}$ marked as adjusted is based on data corrected for baseline (i.e., pre-treatment) differences in the target academically-focused emotion.

3.2.2; R Core Team, 2015). Prior to analysis, *r*s were converted to Fisher's *z* to stabilize their variance and normalize the sampling distribution. For presentation of results, values were reconverted to *r*.

Since all but one of the emotion-correlate models included multiple dependent effect sizes per study, mean weighted correlations were computed using correlated-effects robust variance estimation models (RVE; Hedges, Tipton, & Johnson, 2010; see Supplementary Material for details) provided in the Robumeta-package (Fisher & Tipton, 2015). As the small-sample RVE correction factor implemented in Robumeta requires a minimum of five independent studies for reliable estimation, correlations were aggregated for relations that met this precondition (Tanner-Smith, Tipton, & Polanin, 2016; Tipton, 2015). Additionally, as recommended by Tanner-Smith et al. (2016), we checked our data for outliers based on Tukey's interquartile range (IQR) criterion and re-ran mean effect computations after outlier removal to improve the estimation of degrees of freedom. Similar approaches have been taken by other recently published meta-analyses in educational research and related fields that employed robust variance estimation techniques based on comparable or even smaller primary study samples (e.g., see Friese, Frankenbach, Job, & Loschelder, 2017; Gardella, Fisher, & Teurbe-Tolon, 2017). To generate a more complete picture of available evidence, we provide descriptive information for relations for which at least two effect sizes were coded from the 149 initially reviewed studies.

Based on these preemptive steps, 28 out of the 35 mean effects models (80%) computed for this review were based on reliable estimation. The remaining seven models did not reach the critical threshold of four degrees of freedom required for reliable significance testing of mean effects (Tanner-Smith et al., 2016). However, while the obtained significance value of the estimated mean effect may not be reliable, its magnitude can still yield meaningful information. Accordingly, while *p*-values are reported, we primarily focus on effect sizes (i.e., magnitudes and directions of effects) in our discussion of mean effects and moderators. This approach was also taken due to the number of significance tests conducted (Polanin & Pigott, 2014). Additionally, as RVE estimates heterogeneity (τ^2) using a simple method-of-moments estimator that typically overestimates variation (Scammacca, Roberts, & Stuebing, 2014), we cautiously considered τ^2-values $> .05$ potentially indicative of noteworthy heterogeneity as this amounts to an approximate standard deviation of $\pm$ 0.22 units in the effect size *r*, following similar thresholds as proposed by Cohen (1988; see additional details in Methods section in the Supplementary Material).

As emotion-correlate models were each based on a relatively small number of studies, dichotomous (categorical) moderators were entered in separate meta-regressions to safeguard against loss of power. Because imbalanced numbers of studies per moderator category compromise estimation power in RVE by reducing the degrees of freedom (see Tanner-Smith et al., 2016; for details), analyses were conducted only for emotion-correlate models for which binary moderators yielded a minimum of five studies within, and roughly equal numbers of studies between both moderator categories. Specifically, we ran meta-regressions only if the number of studies assigned to one of the categories did not exceed 65% of the total number of studies included in an emotion-correlate model. For example, for the anxiety-control relation based on 24 independent studies after outlier removal (see Table 2), moderator analyses were not conducted if one of the moderator categories contained 15 or more (i.e., approximately 65%) of the primary studies.

With regard to the moderator ‘emotion focus’, we differentiated between technology-versus academically-focused emotions. For comparisons pertaining to relative universality of relations across gender, we sought to compare effect sizes based on predominantly (i.e., > 75%) female samples to those based on mixed gender compositions as these categories were most commonly observed during coding. For comparisons of mean emotion-correlate relations across continents, we sought to contrast studies based in North American versus Asian, European, or Oceanian contexts, or a combined group of non-North American continents, depending on available effect sizes and categorical distribution. Finally, for ‘type of TBLE’, we sought to compare studies examining technology skill acquisition versus studies examining learning with technology (e.g., ITS, non-intelligent programs, virtual realities; see 2.1) as this dichotomous classification was perceived to capture the most extreme differences in TBL setting. For TBLE versus non-TBLE comparisons, we classified studies as examining the use of basic web/computer tools versus more complex ‘standalone’ systems (i.e., non-intelligent software, ITS, LMS/CMS). In line with the aforementioned criteria, we screened emotion-correlate models based on at least 10 of the 149 initially reviewed independent studies for moderator analysis eligibility in terms of providing balanced subgroups.

2.4.2. Assessment of publication bias

Possible influences of publication bias against nonsignificant or adverse findings were graphically examined through contour-enhanced funnel plots in which *r*s (converted to Fisher's *z*) extracted from the 149 initially reviewed studies were plotted against their standard errors, and symmetry of zero-centered funnel plots inspected with the aid of contour lines representing conventional significance levels ($p < .01$, $p < .05$, $p < .10$; Peters, Sutton, Jones, Abrams, & Rushton, 2008). Additionally, we conducted *a priori* weight function modeling (WFM) to examine the robustness of a mean effects when re-weighting observed effect sizes based on weighting schemes that assign higher probabilities to more significant effect sizes (Vevea & Woods, 2005). As WFM has not been developed for RVE meta-analysis, we created new datasets containing one weighted mean effect size and sampling variance per study (following Borenstein, Hedges, Higgins, & Rothstein, 2009; see Supplementary Material for details). As bias detection requires a reasonable degree of effect size dispersion, analyses were restricted to emotion-correlate relations for which at least 10 independent studies contributed effects (Sterne et al., 2011). The R-packages Weightr (Coburn & Vevea, 2016; weighting schemes provided by; Vevea & Woods, 2005) and Metafor (Viechtbauer, 2010) were used for WFM and contour-enhanced funnel plot production, respectively.

3. Results and discussion

3.1. Frequency of studies across years

Of the final sample of 149 studies included in the initial review and published until May 2014, nearly half of the studies (43%) were published in the first half of the current decade. Delineating the distribution of these 149 studies, Fig. 1 shows that research on emotions in TBLEs has grown tremendously over the past years. Our supplemental literature search for research on emotions in TBLEs yielded 37 additional studies published between June 2014 and July 2018. Full references of all 186 studies are provided in the Supplementary Materials.

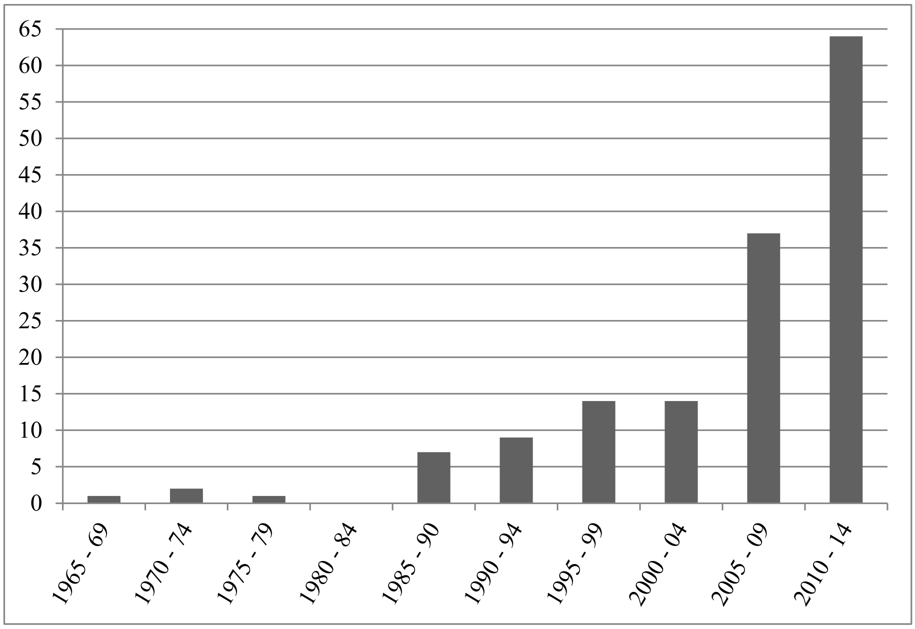

Fig. 1. Number of studies included in the initial review (stage 1) by time period.

3.2. Types of research

3.2.1. Samples and environments

The 149 studies included in our initial review contributed 4.81 effect sizes on average and totaled 19,288 participants. In 48.3% of the studies, the proportion of female participants was between 0.25 and 0.75 (see Table S5 for a summary of study characteristics). Most studies examined emotions of higher education students and adult learners (73.2%). Studies were also mostly conducted in North American contexts (61.1%). Learning/content management systems, non-intelligent learning programs, and ITS were the most frequently studied TBLEs (26.2%, 23.5%, and 21.5%, respectively; for studies that included multiple TBLEs representing different categories, each type of TBLE was coded as a separate entry for this descriptive assessment). TBL typically involved multiple sessions (63.1% of studies) and lasted for more than one hour (59.7%); single sessions (37.6%) and TBL sessions lasting one hour or less (20.8%) were less common.

Our supplemental review of research published between June 2014 and July 2018 revealed similar trends. Specifically, the majority of studies examined emotions of higher education students and adult learners (75.7%) in TBLEs involving non-intelligent learning programs or learning/content management systems (48.6% and 32.4%, respectively). Again, TBL typically involved multiple sessions and lasted for more than one hour (70.3% for both descriptors). More than one third of the studies were conducted in European contexts (37.8%), and 29.7% of studies in North American contexts.

3.2.2. Emotions

The 149 studies included in our initial review yielded 231 emotion assessments and six experimental inductions (i.e., 237 unique examinations of emotions; see Table S5). Paralleling trends in educational emotion research more generally (Pekrun, Goetz, Titz, & Perry, 2002), anxiety was the most commonly investigated emotion (46.6%), followed by enjoyment (24.1%). Academic emotions and technology emotions made up 37.1% and 25.7% of all emotion examinations. Additionally, more than half of the affect assessments measured learners' habitual emotions by means of self-report (59.6%; $N = 231$), while another 35.5% examined self-reported state-level emotions. External observation (4.8%) was relatively rare. For a significant portion (27.7%) of multi-item emotion measures implemented, sample-specific reliabilities were not reported; additionally, single-item assessment, which may differ considerably from multi-item approaches in terms of lower measurement accuracy (e.g., Gogol et al., 2014), was also prevalent. As reliability can impact estimates of effect size (Schmidt & Hunter, 2015), differences in reliability are important to consider when evaluating heterogeneity.

Similar patterns emerged for the 75 emotion assessments considered in our supplemental review of 37 studies published after May 2014. Anxiety again constituted the most frequently examined emotion (54.7%), followed by enjoyment (24.0%) and boredom (14.7%). Examinations of anger/frustration (5.3%) and curiosity/interest (1.3%) were rare, while confusion was not examined by any of the studies. More than half of the assessments involved academic emotions (53.3%). All 75 assessments were based on learners' self-report, and 84.0% of these assessments involved trait-level emotions. Sample-specific reliabilities for multi-item measures were lacking for more than half of the emotion assessments (52.9%).

3.2.3. Correlates

Given the broad range of emotion correlates covered in this synthesis, details on the operationalization of correlates are provided in the Supplementary Material (Tables S4 and S8). Of those correlates considered, control-value appraisals and learning outcomes yielded most of the effect sizes examined in our initial review covering 149 studies (17.8% and 20.4% of the total number of coded effect sizes across all emotions and correlates, $N = 714$; see Table S6). Similarly, of the 224 effect sizes extracted from the 37 studies published after May 2014, most relations pertained to control-value appraisal antecedents and learning outcomes of emotions (56.7% and 17.9%, respectively).

3.3. Relations with antecedents and outcomes

Mean weighted emotion-correlate relations ($\bar{r}$) for relations examined by at least five independent studies included in our initial review are presented in Table 2, and effect sizes for relations examined by less than five independent studies in Table S6. For most relations, enjoyment and anxiety were the only emotions for which sufficient numbers of studies offered effect sizes for statistical integration. Overall, mean weighted effects ranged from small to large ($|.04 \leq \bar{r} \leq .56|$; unless otherwise indicated below, mean weighted effects were significant at $p \leq .05$), with most pronounced effects observed for relations between emotions and control-value appraisals, prior knowledge, and perceptions of overall TBLE quality.

Similarly, observed (i.e., non-weighted) effect sizes extracted as part of our supplemental review ranged from small to large ($|0.00 \leq \bar{r} \leq 0.80|$), with most pronounced effects again emerging for relations between emotions and control-value appraisals. As noted above, confusion was not examined by any of the additional studies included in the supplemental review, such that no additional effect size data was available for correlates of confusion. For comparative purposes, supplemental review findings addressing relations for which mean weighted effects were computed in our initial review are presented below, and those addressing less frequently examined relations in Supplementary Materials.

3.3.1. Antecedents

Corroborating our hypotheses, in our initial review, mean weighted correlations between enjoyment and perceived control as well as positive valuation of learning or the technology at hand were positive and fairly large ($\bar{r} = 0.50$ and 0.56, respectively). In contrast, anxiety was negatively related to both appraisals ($\bar{r} = -0.27$ and -0.13, respectively). Similar emotion-appraisal patterns were observed in studies published after May 2014. Enjoyment was generally positively related to perceived control ($rs = 0.12$–0.68, number of effect sizes [n_{ES}] = 21) and positive valuation of learning/technology ($rs = -0.06$–0.80, $n_{ES} = 13$; note that $r = -0.06$ was the only negative effect size for this relation). Similarly, for anxiety, relations with perceived control were typically negative ($rs = -0.75$ to -0.01, $n_{ES} = 33$), and relations with positive valuation of learning/technology ranged from -0.56 to 0.27 ($n_{ES} = 17$). As such, ranges of observed effects were similar across the initial and supplemental reviews (see Table 2).

In terms of learner differences, in the initial review mean weighted relations between emotions and gender were weak ($\bar{r} = -0.05$ for enjoyment, p-value not estimated, see 3.4.1; $\bar{r} = -0.04$ for anxiety, *n.s.*) implying that male and female learners do not differ in their academically-related or technology-related enjoyment and anxiety within TBL settings. In the studies included in the supplemental review, relations between enjoyment and gender ($r = -0.15$, $n_{ES} = 1$) and anxiety and gender ($r = -0.01$–0.27, $n_{ES} = 2$) were rarely examined but fell into the ranges of effect sizes observed in our initial review (see Table 2). In contrast, mean weighted relations with prior knowledge were more pronounced and, as expected, positive for enjoyment ($\bar{r} = 0.20$), and negative for anxiety ($\bar{r} = -0.28$) in our initial review. Among the studies included in the supplemental review, relations between emotions and prior knowledge were very rarely examined, with one study reporting a positive but weak correlation for enjoyment ($r = 0.08$, $n_{ES} = 1$) again falling within the range of observed effects in our initial review (see Table 2).

Findings also supported several predictions regarding functions of TBLE characteristics for learners' emotions, implying that the CVT's assumptions regarding linkages between cognitive-motivational qualities of learning environments and learners' emotions translate to TBL. For positive emotions, in the initial review we found that learners' enjoyment was positively related to (meta-)cognitive support ($\bar{r} = 0.20$) as well as their evaluations of the overall quality of the TBLE ($\bar{r} = 0.52$). Curiosity/interest also related positively to aesthetic design enhancements, including usage of warm, bright colors (versus greyscale), additions of visual illustrations, or animation, with a small mean weighted effect of $\bar{r} = 0.17$. In our supplemental review, one additional study examined effects for enhanced aesthetic design on learners' curiosity/interest. Observed relations were close to null ($r = -0.02$–0.04, $n_{ES} = 2$) but again fell within the range of observed effects in our initial review (see Table 2). Our supplemental review did not yield any additional findings for relations of enjoyment with (meta-)cognitive support or overall quality of the TBLE.

Mean relations between TBLE characteristics and negative emotions also aligned with our hypotheses. In our initial review, anxiety was negatively related to perceptions of, or experimentally manipulated positive values of learning, achieving, or technology use communicated by peers or instructors ($\bar{r} = -0.16$), as well as to perceived overall TBLE quality ($\bar{r} = -0.17$, p-value not estimated). Similarly, in our supplemental review, one additional study reported a negative but weak association between anxiety and perceived peer/instructor esteem ($r = -.05$) that again fell within the range of observed effects from our initial review (see Table 2).

Furthermore, in our initial review we found that confusion was negatively related to (meta-) cognitive support ($\bar{r} = -0.19$, p-value not estimated), as hypothesized. Surprisingly, however, it was only weakly connected to cognitive conflict ($\bar{r} = .07$, p-value not estimated). Here, we had anticipated a more pronounced positive relation. One implication of this finding is the need to examine more closely under which conditions cognitive conflicts are actually induced. Students may not always detect contradictions or false feedback where presented, and subsequently 'fail' to feel confused. Additionally, by examining learning performance on embedded post-contradiction confusion questions, D'Mello and Graesser (2014) showed that confusion may be higher than students' self-reports reflect. Confusion may imply personal failure to understand or master learning material to some learners which they may not be willing to share. All correlations between confusion and cognitive conflict included in our analysis were based on self-report data, which may help explain the low overall relation obtained. As noted above, because none of the studies included in our supplemental review examined confusion, no additional data was available for these relations.

In addition, several studies included in our initial review examined differences in learner emotions in TBLE versus non-TBLE settings, yielding a weak overall effect on anxiety ($\bar{r} = -.07$, *n.s.*; when adjusting for interindividual differences in pre-treatment emotions, $\bar{r} = -0.13$, *n.s.*), and a small positive effect on enjoyment ($\bar{r} = 0.16$, *n.s.*). These mean effects point to slightly more enjoyment and less anxiety in TBLEs as compared with non-TBLEs, but effects seem to have varied across the different learning settings examined (see 4.4). Interestingly, most studies on anxiety focused on domains like math and statistics, examining learners' course-, learning- or specifically math- (i.e., domain-) related anxiety, while studies targeting learning- or domain-specific enjoyment also included technology-supported reading and writing practice.

Several studies included in our supplemental review also examined differences in learners' anxiety in TBLEs versus non-TBLEs, generally pointing to less anxiety in TBLEs ($r = -0.36$ to -0.08, $n_{ES} = 5$). One study provided an effect size adjusted for pre-treatment differences in emotions, indicating a slightly more pronounced difference ($r = -0.12$). These findings are consistent with the effect sizes observed in our initial review (see also Table 2).

3.3.2. Outcomes

Mean relations between emotions and parameters of the learning process (i.e., engagement, disengagement, strategy use) were also largely in line with our hypotheses. In our initial review, enjoyment was positively related to both engagement and strategy use, with both pooled effects reaching moderate magnitudes ($\bar{r} = 0.45$ and 0.31, respectively), and, as expected, negatively associated with disengagement, although the mean correlation was noticeably smaller ($\bar{r} = 0.10$; p-value not estimated). In line with these patterns, enjoyment was also positively correlated with learning outcomes ($\bar{r} = .18$), to a similar degree as curiosity/interest (small to moderate pooled correlation of 0.20; p-value not estimated). These values are also similar to those observed in prior classroom research on emotions and achievement (see review by Goetz & Hall, 2013). Furthermore, these patterns were largely corroborated by our supplemental review. Specifically, relations of enjoyment with engagement and learning outcomes were again mostly positive and well aligned with the effect sizes observed in our initial review ($rs = 0.05–0.62$, $n_{ES} = 6$, and $rs = -0.06–0.27$, $n_{ES} = 5$, for engagement and learning outcomes respectively; see Table 2 for effect sizes from our initial review).

Anxiety, in contrast, was negatively related to engagement, strategy use, and learning outcomes ($\bar{r} = -0.12$, -0.17, and -0.14, respectively) in our initial review. While these relations are small in magnitude, their directions are in line with prior classroom research (Goetz & Hall, 2013; Hembree, 1988; Pekrun et al., 2002; Seipp, 1991) as well as our predictions. In the supplemental review, observed relations between anxiety and engagement ranged from -0.48 to 0.20 ($n_{ES} = 21$), and relations between anxiety and strategy use ranged from -0.16 to -0.29 ($n_{ES} = 2$), corroborating the hypothesis that anxiety can variable motivational and behavioral effects. Similarly, for anxiety and learning outcomes, effects ranged from -0.54 to 0.34 ($n_{ES} = 30$), although most correlations were negative ($n_{ES} = 25$), consistent with our initial review (see Table 2).

For anger/frustration, confusion, and boredom, mean relations with learning outcomes computed in our initial review were negative as well, but non-significant and even lower than for anxiety ($\bar{r} = -0.07$, -0.09, and -0.08, respectively; Table 2). For negative activating anger/frustration and confusion, variable effects were considered possible (see also 4.4). However, we expected a more pronounced negative relation between boredom and learning outcomes. In the supplemental review, we observed negative relations for anger/frustration and learning outcomes ($r = -0.17$ to -0.06, $n_{ES} = 2$) as well as for boredom and learning outcomes ($r = -0.19$ to -0.12, $n_{ES} = 2$). Interestingly, Tze et al. (2016) also reported a small mean correlation of -0.16 between boredom and achievement in a meta-analysis primarily covering non-TBLEs. Schukajlow and Rakoczy (2016) contend that research on academic outcomes of boredom has typically focused on course-level grades or GPAs rather than test performance. Measures of the former kind often incorporate factors like classroom participation, potentially allowing for more pronounced effects of learning-related boredom on these achievement indicators. Only two of the eleven studies examining relations between boredom and learning outcomes that were included in our review provided correlations based on course grades, while the rest were based on test scores. As such, our findings may not be a product of the technology-based nature of the learning environments examined, but may be due to factors that are just as relevant to non-TBL. Moreover, as noted by Goetz and Hall (2013), even small effects of emotions on achievement can have a strong cumulative impact on long-term learning outcomes.

In conjunction with our findings regarding antecedents of emotions, the pooled effects for outcomes of emotions in both the initial and the supplemental review are largely in line with the idea that TBL and non-TBL are 'not that different in theory' (Daniels & Stupnisky, 2012) in terms of the functional mechanisms of emotions.

3.4. Heterogeneity and moderator analyses

We used the data from our initial review to examine effect size heterogeneity and possible moderators. Between-study effect size heterogeneity τ^2 exceeded .05 in only five of all 35 meta-regression models (Table 2), including pooled correlations between enjoyment and value, engagement, and TBLE-non-TBLE comparisons, between anxiety and TBLE-non-TBLE comparisons (adjusted for baseline anxiety ratings), and between confusion and learning outcomes. None of these models qualified for moderator analyses to examine potential sources of heterogeneity based on the criteria outlined in 3.4.1. However, for three of these models, variation was actually expected. This includes comparisons of enjoyment and anxiety levels in TBLE versus non-TBLE settings, in which the specific designs and thus cognitive-motivational quality of both TBLEs and non-TBLEs may vary across studies, thus impacting learners' emotions in different ways. Furthermore, regarding heterogeneity of relations between confusion and learning outcomes, variation in linkages with achievement outcomes are less surprising considering that 'productive' levels of confusion may foster deep learning (D'Mello, Lehman, Pekrun, & Graesser, 2014), (see Table 1). High τ^2-values for enjoyment-value and enjoyment-engagement were less expected, but the effect size ranges reported in Table 2 indicate that for both models, all correlations were positive. As such, while the magnitude of relations varied across studies, the direction of relations was consistent with our hypotheses.

For emotion-correlate models based on at least 10 independent studies, we examined whether emotion focus, sample gender composition, cultural background (continent), or type of TBLE help explain effect size dispersions. Again, to preserve estimation accuracy, moderator analyses were conducted only if dichotomous categorizations yielded balanced subgroups (see 3.4.1). Table 3 provides the results of the moderator analyses conducted which indicate that overall, mean effects were fairly robust to the moderators examined. The null hypotheses stating that there are no differences between subgroups could be rejected at $p < .05$ in only two of the 18 analyses conducted, and changes in effect size magnitudes in terms of % of variance explained were generally fairly small (see discussion below).

Further supporting the assumption of relative universality, subgroup effects differed only in magnitude, not direction. The only exception to this pattern was the change in direction in anxiety-gender relations when comparing academic with technology-related anxiety (i.e., emotion focus moderation), but the overall effect size remained small in both subgroups. More specifically, correlations between anxiety and perceived control, value, as well as engagement were slightly stronger when anxiety was directed towards technology, but pooled effects remained negative across subgroups, and differences between subgroups amounted to less than 5% of variance explained for the anxiety-control, anxiety-value, and anxiety-engagement relations. Moreover, the difference was only significant for anxiety-control relations.

Relations between technology-related versus academic anxiety and learning outcomes were even more similar ($\bar{r} = -0.15$ and -0.19, respectively). For the aforementioned anxiety-gender relation, subgroup effects point to a weak positive relation between gender and technology anxiety, with females reporting slightly higher levels of anxiety ($\bar{r} = .09$), whereas the relation was negative for academic anxiety, implying lower levels of anxiety for females ($\bar{r} = -0.05$). However, these gender differences in anxiety remained consistently weak even when considering this moderator. Examining moderating effects of subject domains may be more telling for this variable, given that gender differences in emotions have been linked to cultural stereotypes regarding this variable (e.g., Frenzel, Pekrun, & Goetz, 2007a). We were unable to test this assumption due to insufficient numbers of studies for different domains.

Examining potential moderating effects of sample gender composition was only possible for the anxiety-strategy use relation due to the fact that in most studies included in our initial review, gender composition was 'mixed' (i.e., neither male nor female dominant). As suggested by the CVT, relations were virtually equivalent across subgroups.

Regarding subgroup differences by continent as a proxy for cultural background, relations of enjoyment and anxiety with appraisals were consistent in terms of direction across continental comparisons (see Table 2). While the anxiety-value relation appears to be slightly weaker in Asian versus North-American contexts (2.56% difference in terms of variance explained), pooled subgroup effects did not differ significantly in magnitude for these relations. Most notable discrepancies emerged for relations between anxiety and strategy use ($p < .05$) as well as engagement ($p < .10$). In both cases, negative relations were weaker in North American as compared with Asian or multi-continental contexts ('other'), with subgroup differences amounting to 4.84% and 3.24% in variance explained, respectively. These patterns imply that anxiety may have more motivationally and behaviorally debilitating effects in non-North American contexts; however, relations of both enjoyment and anxiety with learning outcomes were practically equivalent across

Table 3
Moderator analyses for select relations conducted in the initial review (stage 1).

Moderator	Relation		95% CI		Moderator subgroup 1			95% CI		Moderator subgroup 2			95% CI	
		b	LL	UL	category	k (n_{ES})	$\bar{r}$	LL	UL	category	k (n_{ES})	$\bar{r}$	LL	UL
Emotion focus	anxiety-control	.20*	.01	.38	technology	12 (23)	-.36***	-.49	-.22	academic	11 (16)	-.17*	-.31	-.03
	anxiety-value	.13	-.10	.37	technology	8 (10)	-.21*	-.38	-.04	academic	8 (13)	-.08	-.27	.11
	anxiety-gender	.10	-.08	.28	technology	8 (17)	.09	-.03	.20	academic	8 (8)	-.05	-.18	.15
	anxiety-engagement	.11	-.12	.34	technology	7 (12)	-.18*	-.34	-.00	academic	5 (12)	-.06	-.28	.16
	anxiety-learning outcomes	-.04	-.17	.09	technology	11(15)	-.15*	-.26	-.04	academic	10 (16)	-.19**	-.28	-.10
Gender	anxiety-strategy use	-.01	-.25	.28	mostly female	6 (9)	-.11	-.32	.11	mixed sample	7 (14)	-.13	-.35	.10
Continent	enjoyment-control	.23	-.06	.49	North America	7 (8)	.43***	.27	.56	Other	5 (7)	.60(†)	.37	.76
	enjoyment-learning outcomes	.05	-.11	.21	North America	5 (7)	.15(†)	.00	.29	Europe	5 (11)	.20(†)	.07	.32
	anxiety-control	.07	-.13	.27	North America	12 (22)	-.30**	-.44	-.13	Asia	11 (18)	-.22**	-.34	-.10
	anxiety-value	.16	-.09	.39	North America	7 (15)	-.23*	-.00	-.43	Asia	9 (12)	-.07	-.22	.09
	anxiety-gender	.04	-.12	.20	North America	11 (18)	.03	-.10	.15	Other	10 (10)	.06	-.06	.18
	anxiety-engagement	-.18†	-.36	.01	North America	7 (20)	-.04	-.22	.14	Other	7 (11)	-.22**	-.33	-.11
	anxiety-strategy use	.22*	.03	.40	North America	7 (11)	-.04	-.14	.06	Asia	7 (14)	-.26*	-.44	-.06
	anxiety-learning outcomes	-.03	-.17	.11	North America	19 (33)	-.13*	-.22	-.04	Other	10 (15)	-.16*	-.27	-.04
Type of TBLE	anxiety-prior knowledge	.04	-.13	.21	technology skills	12 (52)	-.30***	-.40	-.19	learning with technology	10 (16)	-.28**	-.38	-.12
	anxiety-TBLE vs. non-TBLE	.05	-.20	.10	web/comp. tools	5 (7)	-.05(†)	-.09	-.01	TBLE systems	8 (13)	-.10	-.25	.05
	anxiety-learning outcomes	.03	-.09	.16	technology skills	10 (14)	-.16**	-.26	-.07	learning with technology	19 (34)	-.13**	-.23	-.04
		.14†	-.03	.30	LMS/CMS	5(9)	-.13(†)	-.31	-.07	non-intelligent programs	9 (20)	-.05	-.19	.09

Note. LMS/CMS = learning/content management systems. Web/comp. tools = basic web or computer tools such as text annotation software. TBLE systems = combined category of non-intelligent learning software and ITS (i.e., all studies examining "standalone" systems designed to teach specific content). Gender was coded 0 = male 1 = female. 'Mostly female' refers to samples in which the proportion of females was < 0.75). For 'mixed samples', proportion of females was between 0.25 and 0.75). LL = lower limit UL = upper limit of 95% confidence interval (CI). ***$p < .001$. **$p < .01$. *$p < .05$..(†)p-value not reliable ($df < 4$ for t-test of $\bar{r}$).

subgroups, in line with our hypotheses. Moreover, gender differences in anxiety levels did not differ across cultural subgroups ($\bar{r}$ = .03 and .06 for North American and non-North American contexts, respectively).

When examining whether the type of TBLE affected emotion-correlate relations, one marginally significant difference emerged for relations between anxiety and learning outcomes in specifically LMS/CMS-based environments versus learning with other non-intelligent programs. However, both correlations were negative ($\bar{r} = -.13$ and $-.05$, respectively), and the difference amounts to less than 1% of variance explained, implying that the type of TBLE does not profoundly modulate emotion-correlate relations. For the other relations examined (i.e., anxiety-prior knowledge; anxiety-learning outcomes when acquiring technology skills versus learning with technology; anxiety in TBLEs vs. non-TBLEs), subgroup differences were even smaller.

Taken together, the moderator variables considered in the present review did not substantially contribute to explaining heterogeneity. While this is in line with our assumptions regarding the relative universality of cause-effect mechanisms of emotions across learning contexts, it leaves variance to be explained (see section 5).

3.5. Analysis of possible publication bias

3.5.1. Funnel plots

Funnel plots for relations based on at least 10 independent studies included in our initial review are presented in Fig. S1. Among the 11 plots produced, effect size distributions for relations of enjoyment with control- and value appraisals were most suspicious in terms of potential influence of publication bias. Specifically, *r*s for both relations were exclusively positive and significant, implying that suppression of non-significant findings as well as prejudice against negative relations may have been operating. Such patterns may indicate robust effects (S. Kepes, personal communication, November 17, 2016). However, the plots suggest that they could also be due to 'small-study effects', that is, large effects derived from small studies which are potentially larger than those reported by larger studies (Sterne et al., 2011). Such patterns may be the result of bias in that the chance for smaller studies to be published is increased if they are able to provide strong effects. For anxiety-appraisal relations, in contrast, the plots pointed to "missing" positive correlations, with most effects reported being negative, but the observed correlations also included non-significant effects.

For relations involving more distal emotion antecedents, the anxiety-prior knowledge plot revealed mostly negative correlations, but these were relatively evenly distributed in terms of their statistical significance. The plots for anxiety-gender relations and for anxiety in TBLEs vs. non-TBLEs comparisons included mostly small and non-significant positive and negative correlations. Such patterns are less likely to be produced by publication bias.

For relations involving outcome variables, anxiety-engagement, anxiety-strategy use, and anxiety-learning outcomes correlations were clustered on the left-hand side of the funnel (i.e., mostly negative). However, distributions between significant and non-significant findings were relatively balanced, alleviating concerns somewhat. Similarly, the enjoyment-learning outcomes funnel included mostly positive, but a relatively even number of significant and non-significant correlations. The high degree of consistency of effect sizes for emotion-correlate relations across the initial and supplemental reviews (see section 4) suggests that these conclusions are also valid for the studies included in the supplemental review.

3.5.2. Weight function modeling (WFM)

Results of the WFM analyses computed in the initial review are reported in Table S7. Based on the previous funnel plot inspection and following recommendations by Kepes and colleagues (Kepes, Banks, McDaniel, & Whetzel, 2012), WFM was not conducted for the enjoyment-control and -value relations (based exclusively on significant correlations) as WFM can yield nonsensical results when there is little or no variation in the *p*-values of effect sizes. Moreover, due to the small number of studies for the enjoyment-learning outcomes relation, WFM did not converge (i.e., no estimates were produced) and was thus excluded from the analysis.

WFM estimates the potential influence of bias against non-significant findings may on meta-analytic results and should be interpreted in terms of the degree of change between the originally computed and the adjusted mean effect (Kepes et al., 2012). For the eight remaining relations based on 10 or more independent studies eligible for WFM (i.e., relations of anxiety with appraisals, gender, prior knowledge, engagement, strategy use, learning outcomes, and TBLE vs. non-TBLE comparisons), funnel plots revealed relatively balanced distributions of significant and non-significant relations, such that little impact of re-weighting effect sizes based on significance levels may be expected. Accordingly, with the exception of the severe one-tailed selection function, mean correlations remained fairly stable under the different weighting conditions. More specifically, across moderate one-tailed, moderate two-tailed, and severe two-tailed weighting models, $|\Delta\bar{r}s|$ between the original and the adjusted mean effects for the eight aforementioned relations ranged from 0.01 to 0.04, indicating that publication bias could be present under these conditions, although the effects would likely be only small to moderate, and conclusions would not be substantially altered (Kepes, Banks, & Oh, 2014).

Under the severe one-tailed selection condition, the picture changed more drastically. Here, $|\Delta\ \bar{r}\ s|$ range from 0.05 to 0.13 across the eight relations examined, implying that bias could be severe. The most extreme change observed for the anxiety-gender relation, where the originally small positive effect ($\bar{r}$ .05) was adjusted to a small negative correlation $\bar{r} = -.08$). This may overturn our original conclusion in terms of the direction of the effect, but the adjusted $\bar{r}$ remains fairly weak in terms of magnitude. For the seven remaining relations, again, meta-analytic conclusions would be consistent with our current findings in terms of both magnitude and direction of effect, given that none of the adjusted means lead to changes in effect size categorization as small, medium, or large by conventional standards (Cohen, 1988). In other words, even under severe one-tailed bias conditions, the bias-adjusted estimates and actual pooled effects yielded comparable estimates in terms of the direction and size of effects. It is also important to note that bias-adjusted effects are not to be interpreted as a better or the 'true' effect as they are based on a hypothetical rearrangement of the observed data (Kepes et al., 2012).

4. Conclusions

Understanding and supporting emotional processes of TBL has become a paramount goal across different research communities. This two-stage review sought to contribute to the growing and increasingly important field of affective TBL research by integrating extant evidence for causes and effects of emotions across a broad variety of technologies using an established framework in educational emotion research, the CVT. The goal is to facilitate moving towards a more unified perspective of emotional underpinnings of learning not only with technology, but learning more generally.

Overall, the patterns of effect size computed in our initial review, which covered 149 studies published until May 2014, largely support our hypotheses and thus attest to the applicability of the CVT to TBLEs. Moreover, the findings from our supplemental review of 37 studies published more recently (i.e., between June 2014 and July 2018) provide additional support for the emotion-correlate relations considered and suggest similar patterns of relations in terms of both magnitude and direction of effects. Overall, the consistency of the results across the two reviews is remarkable and attests to the robustness of the findings.

From a CVT perspective, the findings are encouraging. They imply that the theory offers a foundational framework for fundamental mechanisms of emotions across diverse learning environments and can inform their emotionally sound design (Astleitner & Leutner, 2014; Lester et al., 2014). At a very basic level, this includes fostering optimal levels of subjective control and value of learning tasks to promote enjoyment and curiosity, and reduce anxiety. One way to achieve the former, as our data suggest, is by designing aesthetically appealing environments, and providing cognitive and metacognitive scaffolding to support the learning process. On the other hand, the findings suggest that negative emotions like confusion, but potentially also anger or boredom, can be beneficial to learning under certain circumstances, likely only to the degree that they promote deeper engagement with contents and can be successfully resolved.

The value of these insights notwithstanding, they are subject to certain limitations that provide important directions for future research. First, our analysis was exclusively based on published studies, and was only able to examine potential sensitivity to bias for a subset of the relations tested. However, by searching edited conference proceedings for inclusion in our initial review, we made an attempt to include "grey literature" that is typically subject to less rigorous review processes and more likely to include non-significant or adverse findings (Kepes et al., 2012). When interpreting the results of our analyses for publication bias, it is important to consider that they provide indirect evidence and do not allow firm conclusions (Baldwin, Christian, Berkeljon, Shadish, & Bean, 2012, p. 285). Fortunately, the WFM analyses suggest little impact on the overall conclusions of the meta-analytic findings supplementing this review.

Second, although the present review includes data from 186 studies, the number of studies and available effect sizes per emotion-correlate relation was relatively small for some of the emotion-correlate relations, especially for relations involving emotions other than enjoyment and anxiety. This was true for both the 149 studies included in our main review as well as the 37 additional studies surveyed in our supplemental review of recently published research. One issue closely related to this limitation pertains to the unexplained heterogeneity observed in effect size distributions computed as part of our initial review. Given the methodological diversity of the primary studies considered, heterogeneity was anticipated and handled through the *a priori* decision to compute random-effects models, and moderators examined for models based on sufficient numbers of primary studies. We also checked for possibilities to investigate potential influences of several methodological moderators discussed below in addition to those examined in the review (i.e., gender composition; cultural context; object focus of emotion; type of TBLE). However, due to the relatively small number of studies for most meta-analytic models and small sample sizes for moderator subgroups compromising accuracy of estimating moderator effects, we were unable to examine these characteristics statistically.

One factor that likely contributed to variability is lack of psychometric equivalence of measures across studies. As noted, a substantial number of studies in both the initial and the supplemental review did not provide information about reliabilities, which precludes correcting for measurement accuracy (Schmidt & Hunter, 2015) and examining whether differences in reliability account for heterogeneity. Additionally, although studies were required to provide sample items, descriptions of coding practices, or experimental protocols, it cannot be ruled out that construct operationalizations differed in terms of validity. More fine-grained analyses of moderator subcategories, considering state-versus trait-level assessments of emotions and correlate constructs (Bieg et al., 2013), and greater within-construct differentiation (e.g., extrinsic vs. intrinsic motivation; anger vs. frustration; relative effectiveness of TBLE design in enhancing aesthetic appeal or providing support) may also help explain effect size variation.

However, while it must be acknowledged that limitations in terms of sample size precludes analyzing all of the emotion-correlate relations that may be of interest from a theory perspective, we believe taking stock of available evidence as done in the present review to be a useful endeavor even if the number of primary studies is still somewhat limited for some emotions. Specifically, the present review provides cumulative evidence for antecedents and effects of two of the most frequently occurring and important emotions in terms of impact on learning and wellbeing, namely anxiety and enjoyment. Additionally, systematizing extant evidence allows for identifying open questions and directions for future work. This may be particularly valuable for emergent research fields such the one addressed in this study.

As such, our findings document cumulative evidence for enjoyment and anxiety, but also imply that there is a clear need for more systematic research on other emotions. Similarly, we need more research targeting the appraisals underlying emotions in TBLEs and how these appraisals can be impacted through their design. For instance, our data suggest that not all of the approaches aimed at providing (meta-) cognitive support to foster enjoyment were equally able to do so. To the degree that advanced TBLEs are able to adapt to individual learners, differences in needing cognitive support can be taken into account, as these differences likely impact learners' emotions. Related to this issue, our analysis points to a dearth of studies on how autonomy support can influences learners' emotions in TBLEs, another important pillar in the CVT framework that pertains to both the cognitive and the motivational quality of learning environments (Pekrun & Perry, 2014). Providing autonomy support typically entails giving individuals choices in regulating their own learning process, which implies a need for balancing degrees of freedom and sufficient guidance to keep enjoyment and curiosity high.

On a related note, the present review also revealed that possible nonlinear relations between certain emotions and correlates are still rarely considered, thus precluding meta-analytic examinations of nonlinear dynamics. For example, as noted in Section 1.1.1, the relation between perceived control and boredom may be curvilinear in nature, such that boredom may arise either when control is too high or too low (Acee et al., 2010; Goetz & Hall, 2014). Similarly, the relations between some emotions, such as anxiety, and learning outcomes could be curvilinear. Given the lack of evidence on nonlinear relations, the present review revealed linear components of emotion-correlate relations, but could not examine their possible nonlinear components. Future research needs to attend to such relations and produce cumulative findings that can be meaningfully integrated for informing evidence-based practice (e.g., in terms of designing TBL interventions that foster optimal levels of challenge and perceived control).

In sum, the present review indicates that emotions are important drivers of learning in technology-based settings, and that learners' emotional experiences can be shaped by the characteristics of those settings. Furthermore, this review suggests that the CVT can serve as a meta-theoretical frame for guiding future research efforts in these directions, and provides a basis for meaningfully consolidating empirical findings to support evidence-based practice as research on emotions in TBLEs further accumulates.

References

Acee, T. W., Kim, H., Kim, H. J., Kim, J., Hsiang-Ning, R. C., Kim, M., et al. The Boredom Research Group. (2010). Academic boredom in under- and overchallenging situations. *Contemporary Educational Psychology, 35*, 17–27. https://doi.org/10.1016/j.cedpsych.2009.08.002.

Ahmed, W., van der Werf, G., Kuyper, H., & Minnaert, A. (2013). Emotions, self-regulated learning, and achievement in mathematics: A growth-curve analysis. *Journal of Educational Psychology, 105*, 150–161. https://doi.org/10.1037/a0030160.

Ainley, M., & Hidi, S. (2014). Interest and enjoyment. In R. Pekrun, & L. Linnenbrink-Garcia (Eds.). *International handbook of emotions in education* (pp. 205–227). New York, NY: Taylor & Francis.

Artino, A. R. J. (2009). Think, feel, act. *Journal of Computing in Higher Education, 21*, 146–166. https://doi.org/10.1007/s12528-009-9020-9.

Astleitner, H., & Leutner, D. (2014). Designing instructional technology from an emotional perspective. *Journal of Research on Computing in Education, 32*, 497–510. https://doi.org/10.1080/08886504.2000.10782294.

Azevedo, R., Johnson, A., Chauncey, A., & Burkett, C. (2010). Self-regulated learning with MetaTutor: Advancing the science of learning with MetaCognitive tools. In M. S. Khine, & I. M. Saleh (Eds.). *New science of learning (Part II)* (pp. 225–247). New York, NY: Springer.

Atkinson, J., & Feather, N. (1966). *A theory of achievement motivation.* New York, NY: Wiley and Sons.

Baldwin, S. A., Christian, S., Berkeljon, A., Shadish, W. R., & Bean, R. (2012). The effects of family therapies for adolescent delinquency and substance abuse: A meta-analysis. *Journal of Marital and Family Therapy, 38*, 281–304 10.111/j.1752-0606.2011.00248.x.

Beasley, R. E., & Waugh, M. L. (1995). Cognitive mapping architectures and hypermedia disorientation: An empirical study. *Journal of Educational Multimedia and Hypermedia, 4*, 239–255.

Beilock, S. L., Gunderson, E. A., Ramirez, G., & Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences, 107*, 1860–1863. https://doi.org/10.1073/pnas.

Bieg, M., Goetz, T., & Hubbard, K. (2013). Can I master it and does it matter? An intraindividual analysis on control–value antecedents of trait and state academic emotions. *Learning and Individual Differences, 28*, 102–108. https://doi.org/10.1016/j.lindif.2013.09.006.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis.* Chichester, U.K: John Wiley & Sons.

Butz, N. T., Stupnisky, R. H., & Pekrun, R. (2015). Students' emotions for achievement and technology use in synchronous hybrid graduate programmes. *Research in Learning Technology, 23*, 1629. https://doi.org/10.3402/rlt.v23.26097.

Calvo, R., & D'Mello, S. K. (2011). *New perspectives on affect and learning technologies.* New York, NY: Springer.

Coburn, K. M., & Vevea, J. L. (2016). *Package "weightr": Estimating weight-function models for publication bias in R.* Retrieved from https://cran.r-project.org/web/packages/weightr/index.html.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Covington, M. (1984). The self-worth theory of achievement motivation: Findings and implications. *The Elementary School Journal, 85*, 4–20. https://doi.org/10.1086/461388.

Cromley, J., Azevedo, R., & Olson, E. (2005). Self-regulation of learning with multiple representations in hypermedia. In C.-K. Looi, G. McCalla, B. Bredeweg, & J. Breuker (Eds.). *Artificial intelligence in education* (pp. 184–191). Amsterdam: IOS Press.

D'Mello, S. K. (2013). A selective meta-analysis on the relative incidence of discrete affective states during learning with technology. *Journal of Educational Psychology, 105*, 1082–1099. https://doi.org/10.1037/a0032674.

D'Mello, S. K., & Graesser, A. C. (2014). Confusion. In R. Pekrun, & L. Linnenbrink-Garcia (Eds.). *International handbook of emotions in education* (pp. 289–310). New York, NY: Taylor & Francis.

D'Mello, S. K., Lehman, B., Pekrun, R., & Graesser, A. C. (2014). Confusion can be beneficial for learning. *Learning and Instruction, 29*, 153–170. https://doi.org/10.1016/j.learninstruc.2012.05.003.

Daniels, L. M., & Stupnisky, R. H. (2012). Not that different in theory: Discussing the control-value theory of emotions in online learning environments. *Internet and Higher Education, 15*, 222–226. https://doi.org/10.1016/j.iheduc.2012.04.002.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*, 319–340. https://doi.org/10.2307/249008.

Fisher, Z., & Tipton, E. (2015). *robumeta: Robust variance meta-regression (R package version 1.6).* Retrieved from http://CRAN.R-project.org/package=robumeta.

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist, 56*, 218–226. https://doi.org/10.1037/0003-066X.56.3.218.

Frenzel, A. C., Pekrun, R., & Goetz, T. (2007a). Girls and mathematics —a "hopeless" issue? A control-value approahc to gender differences in emotions towards mathematics. *European Journal of Psychology of Education, 22*, 497–514. https://doi.org/10.1007/BF03173468.

Frenzel, A. C., Pekrun, R., & Goetz, T. (2007b). Perceived learning environment and students' emotional experiences: A multilevel analysis of mathematics classrooms. *Learning and Instruction, 17*, 478–493. https://doi.org/10.1016/j.learninstruc.2007.09.001.

Friese, M., Frankenbach, J., Job, V., & Loschelder, D. D. (2017). Does self-control training improve self-control? A meta-analysis. *Perspectives on Psychological Science, 12*, 1077–1099. https://doi.org/10.1177/1745691617697076.

Gardella, J. H., Fisher, B. W., & Teurbe-Tolon, A. R. (2017). A systematic review and meta-analysis of cyber-victimization and educational outcomes for adolescents. *Review of Educational Research, 87*, 283–308. https://doi.org/10.3102/0034654316689136.

Gaspard, H., Dicke, A.-L., Flunger, B., Schreier, B., Häfner, I., Trautwein, U., et al. (2015). More value through greater differentiation: Gender differences in value beliefs about math. *Journal of Educational Psychology, 107*, 663–677. https://doi.org/10.1037/edu0000003.

Goetz, T., Bieg, M., Lüdtke, O., Pekrun, R., & Hall, N. C. (2013a). Do girls really experience more anxiety in mathematics? *Psychological Science, 24*, 2079–2087. https://doi.org/10.1177/0956797613486989.

Goetz, T., & Hall, N. C. (2013). Emotion and achievement in the classroom. In J. Hattie, & E. M. Anderman (Eds.). *International guide to student achievement* (pp. 192–195). London: Routledge.

Goetz, T., & Hall, N. C. (2014). Academic boredom. In R. Pekrun, & L. Linnenbrink-Garcia (Eds.). *International handbook of emotions in education* (pp. 311–330). New York, NY: Taylor & Francis.

Goetz, T., Lüdtke, O., Nett, U. E., Keller, M. M., & Lipnevich, A. A. (2013b). Characteristics of teaching and students' emotions in the classroom: Investigating differences across domains. *Contemporary Educational Psychology, 38*, 383–394. https://doi.org/10.1016/j.cedpsych.2013.08.001.

Gogol, K., Brunner, M., Goetz, T., Martin, R., Ugen, S., Keller, U., ... Preckel, F. (2014). "My questionnaire is too long!" the assessments of motivational-affective constructs with three-item and single-item measures. *Contemporary Educational Psychology, 39*, 188–205. https://doi.org/10.1016/j.cedpsych.2014.04.002.

Graesser, A. C., Chipman, P., King, B., McDaniel, B., & D'Mello, S. (2007). Emotions and

learning with AutoTutor. *Frontiers in Artificial Intelligence and Applications, 158*, 569–571.
Graesser, A. C., D'Mello, S. K., & Strain, A. C. (2014). Emotions in advanced learning technologies. In R. Pekrun, & L. Linnenbrink-Garcia (Eds.). *International handbook of emotions in education* (pp. 473–493). New York, NY: Taylor & Francis.
Graham, S., & Taylor, A. Z. (2014). An attributional approach to emotional life in the classroom. In R. Pekrun, & L. Linnenbrink-Garcia (Eds.). *International handbook of emotions in education* (pp. 96–119). New York, NY: Taylor & Francis.
Gunn, C., French, S., McLeod, H., McSporran, M., & Conole, G. (2002). Gender issues in computer-supported learning. *Research in Learning Technology, 10*, 32–44. https://doi.org/10.1080/0968776020100106.
Hagtvet, K. A., & Benson, J. (1997). The motive to avoid failure and test anxiety responses: Empirical support for integration of two research traditions. *Anxiety, Stress & Coping, 10*, 35–57. https://doi.org/10.1080/10615809708249294.
Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1*, 39–65. https://doi.org/10.1002/jrsm.5.
Hedges, L. V., & Vevea, J. L. (1998). Fixed and random-effects models in meta-analysis. *Psychological Methods, 3*(4), 486–504.
Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety: A meta-analysis. *Review of Educational Research, 58*, 47–77. https://doi.org/10.2307/1170348.
Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*, 111–127. https://doi.org/10.1207/s15326985ep4102_4.
Homer, B. D., Hayward, E. O., Frye, J., & Plass, J. L. (2012). Gender and player characteristics in video game play of preadolescents. *Computers in Human Behavior, 28*, 1782–1789. https://doi.org/10.1016/j.chb.2012.04.018.
Kepes, S., Banks, G. C., McDaniel, M., & Whetzel, D. (2012). Publication bias in the organizational sciences. *Organizational Research Methods, 15*, 624–662. https://doi.org/10.1177/1094428112452760.
Kepes, S., Banks, G. C., & Oh, I.-S. (2014). Avoiding bias in publication bias research: The value of "null" findings. *Journal of Business and Psychology, 29*, 183–203. https://doi.org/10.1007/s10869-012-9279-0.
King, R. B., & Gaerlan, J. M. (2014). High self-control predicts more positive emotions, better engagement, and higher achievement in school. *European Journal of Psychology of Education, 29*, 81–100. https://doi.org/10.1007/s10212-013-0188-z.
Lajoie, S. P., & Azevedo, R. (2006). Teaching and learning in technology-rich environments. In P. A. Alexander, & P. H. Winne (Eds.). *Handbook of educational psychology* (pp. 803–821). (2nd ed.). Mahwah, N.J: Erlbaum.
Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping.* New York, N.Y: Guilford Press.
Lester, J. C., Spires, H. A., Nietfeld, J. L., Minogue, J., Mott, B. W., & Lobene, E. V. (2014). Designing game-based learning environments for elementary science education: A narrative-centered learning perspective. *Information Sciences, 264*, 4–18. https://doi.org/10.1016/j.ins.2013.09.005.
Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage.
Loderer, K., Pekrun, R., & Plass, J. L. (2018). Emotional foundations of game-based learning. In J. L. Plass, B. D. Homer, & R. E. Mayer (Eds.). *Handbook of game-based learning*Cambridge, MA: MIT Press (in press).
Mandler, G., & Sarason, S. B. (1952). A study of anxiety and learning. *Journal of Abnormal and Social Psychology, 47*, 166–173. https://doi.org/10.1037/h0062855.
Marsh, H. W. (1993). Academic self-concept: Theory, measurement, and research. In J. M. Suls (Vol. Ed.), *The self in social perspective: . Vol. 4. Psychological perspectives on the self* (pp. 59–98). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.
McKeachie, W. J. (1951). Anxiety in the college classroom. *Journal of Educational Research, 45*, 153–160.
McNamara, D. S., Jackson, G. T., & Graesser, A. C. (2010). Intelligent tutoring and games (ITaG). In Y. K. Baek (Ed.). *Gaming for classroom-based learning: Digital role-playing as a motivator of study* (pp. 44–65). Hershey, PA: IGI Global.
Moos, D. C., & Azevedo, R. (2009). Learning with computer-based learning environments: A literature review of computer self-efficacy. *Review of Educational Research, 79*, 576–600. https://doi.org/10.3102/0034654308326083.
Moos, D. C., & Marroquin, E. (2010). Multimedia, hypermedia, and hypertext: Motivation considered and reconsidered. *Computers in Human Behavior, 26*, 265–276. https://doi.org/10.1016/j.chb.2009.11.004.
Moreno, R. (2006). Does the modality principle hold for different media? A test of the method-affects-learning hypothesis. *Journal of Computer Assisted Learning, 22*, 149–158. https://doi.org/10.1111/j.1365-2729.2006.00170.x.
Moreno, R. (2012). Application of technology to learning. In K. R. Harris, S. Graham, T. Urdan, A. G. Bus, S. Major, & H. L. Swanson (Vol. Eds.), *Applications to learning and teaching: . Vol. 3. APA educational psychology handbook* (pp. 427–449). Washington, D. C: APA.
Muis, K. R., Chevrier, M., & Singh, C. A. (2018). *The role of epistemic emotions in personal epistemology and self-regulated learning. Educational Psychologist.* Advance online publicationhttps://doi.org/10.1080/00461520.2017.1421465.
Muis, K. R., Psaradellis, C., Lajoie, S. P., Di Leo, I., & Chevrier, M. (2015). The role of epistemic emotions in mathematics problem solving. *Contemporary Educational Psychology, 42*, 172–185. https://doi.org/10.1016/j.cedpsych.2015.06.003.
Murayama, K., Miyatsu, T., Buchli, D., & Storm, B. C. (2014). Forgetting as a consequence of retrieval: A meta-analytic review of retrieval-induced forgetting. *Psychological Bulletin, 140*, 1383–1409. https://doi.org/10.1037/a0037505.
Niculescu, A. C., Tempelaar, D., Dailey-Hebert, A., Segers, M., & Gijselaers, W. (2015). Exploring the antecedents of learning-related emotions and their relations with achievement outcomes. *Frontline Learning Research, 3*, 1–17. https://doi.org/10.14786/flr.v3i1.136.
Noteborn, G., Bohle Carbonell, K., Dailey-Hebert, A., & Gijselaers, W. (2012). The role of emotions and task significance in Virtual Education. *The Internet and Higher Education, 15*, 176–183. https://doi.org/10.1016/j.iheduc.2012.03.002.
Pekrun, R. (1992). Expectancy-value theory of anxiety: Overview and implications. In D. G. Forgays, T. Sosnowski, & K. Wrzesniewski (Eds.). *Anxiety: Recent developments in self-appraisal, psychophysiological and health research* (pp. 23–41). Washington, DC: Hemisphere.
Pekrun, R. (2005). Progress and open problems in educational emotion research. *Learning and Instruction, 15*, 497–506. https://doi.org/10.1016/j.learninstruc.2005.07.014.
Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*, 315–341. https://doi.org/10.1007/s10648-006-9029-9.
Pekrun, R. (2009). Global and local perspectives on human affect: Implications of the control-value theory of achievement emotions. In M. Wosnitza, S. A. Karabenick, A. Efklides, & P. Nenninger (Eds.). *Contemporary motivation research: From global to local perspectives* (pp. 97–115). Cambridge, MA: Hogrefe.
Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology, 36*, 36–48. https://doi.org/10.1037/t21196-000.
Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement. *Educational Psychologist, 37*, 91–105. https://doi.org/10.1207/S15326985EP3702_4.
Pekrun, R., Hall, N. C., Goetz, T., & Perry, R. P. (2014). Boredom and academic achievement: Testing a model of reciprocal causation. *Journal of Educational Psychology, 106*, 696–710. https://doi.org/10.1037/a0036006.
Pekrun, R., & Perry, R. P. (2014). Control-value theory of achievement emotions. In R. Pekrun, & L. Linnenbrink-Garcia (Eds.). *International handbook of emotions in education* (pp. 120–141). New York, NY: Taylor & Francis.
Pekrun, R., & Stephens, E. J. (2012). Academic emotions. In K. R. Harris, S. Graham, T. Urdan, S. Graham, J. M. Royer, & M. Zeidner (Eds.). *APA educational psychology handbook, Vol 2: Individual differences and cultural and contextual factors* (pp. 3–31). Washington: American Psychological Association.
Pekrun, R., Vogl, E., Muis, K. R., & Sinatra, G. M. (2017). Measuring emotions during epistemic activities: The epistemically-related emotion scales. *Cognition & Emotion.* https://doi.org/10.1080/02699931.2016.1204989.
Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2008). Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology, 61*, 991–996. https://doi.org/10.1016/j.jclinepi.2007.11.010.
Plass, J. L., & Kaplan, U. (2016). Emotional design in digital media for learning. In S. Y. Tettegah, & M. P. McCreery (Eds.). *Emotions, technology, and learning* (pp. 131–161). London: Academic Press.
Polanin, J. R., & Pigott, T. D. (2014). The use of meta-analytic statistical significance testing. *Research Synthesis Methods, 6*, 63–73. https://doi.org/10.1002/jrsm.1124.
Putwain, D. W. (2008). Deconstructing test anxiety. *Emotional & Behavioural Difficulties, 13*, 141–155. https://doi.org/10.1080/13632750802027713.
Putwain, D. W., Sander, P., & Larkin, D. (2013). Academic self-efficacy in study-related skills and behaviours: Relations with learning-related emotions and academic success. *British Journal of Educational Psychology, 83*, 633–650. https://doi.org/10.1111/j.2044-8279.2012.02084.x.
R Core Team (2015). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/.
Ranellucci, J., Hall, N. C., & Goetz, T. (2015). Achievement goals, emotions, learning, and performance: A process model. *Motivation Science, 1*, 98–120. https://doi.org/10.1037/mot0000014.
Sabourin, J. L., & Lester, J. C. (2014). Affect and engagement in game-based learning environments. *IEEE Transactions on Affective Computing, 5*, 45–56. https://doi.org/10.1109/T-AFFC.2013.27.
Scammacca, N., Roberts, G., & Stuebing, K. K. (2014). Meta-analysis with complex research designs: Dealing with dependence from multiple measures and multiple group comparisons. *Review of Educational Research, 84*, 328–364. https://doi.org/10.3102/0034654313500826.
Schmidt, F. L., & Hunter, J. E. (2015). *Methods of meta-analysis: Correcting error and bias in research findings* (3rd ed.). Los Angeles, CA: Sage.
Schukajlow, S., & Rakoczy, K. (2016). The power of emotions: Can enjoyment and boredom explain the impact of individual preconditions and teaching methods on interest and performance in mathematics? *Learning and Instruction, 44*, 117–127. https://doi.org/10.1016/j.learninstruc.2016.05.001.
Seipp, B. (1991). Anxiety and academic performance: A meta-analysis of findings. *Anxiety Research, 4*, 27–41. https://doi.org/10.1080/08917779108248762.
Seipp, B., & Schwarzer, C. (1996). Cross-cultural anxiety research: A review. In C. Schwarzer, & M. Zeidner (Eds.). *Stress, anxiety, and coping academic settings* (pp. 13–68). Tuebingen: Francke.
Shuman, V., & Scherer, K. R. (2014). Concepts and Structures of Emotions. In R. Pekrun, & L. Linnenbrink-Garcia (Eds.). *International Handbook of Emotions in Education* (pp. 13–35). New York, N.Y.; London: Taylor and Francis.
Sieber, J. E., O'Neil, H. F., & Tobias, S. (1977). *Anxiety, learning, and instruction.* New York, NY: Routledge.
Silvia, P. J. (2008). Interest - the curious emotion. *Current Directions in Psychological Science, 17*(1), 57–60. https://doi.org/10.1111/j.1467-8721.2008.00548.x.
Spielberger, C. D., & Vagg, P. R. (1995). Test anxiety: A transactional process. In C. D. Spielberger, & P. R. Vagg (Eds.). *Test anxiety: Theory, assessment, and treatment* (pp. 3–14). Washington, DC: Taylor & Francis.
Stengel, E. (1936). Prüfungsangst und Prüfungsneurose. *Zeitschrift für Psychoanalytische Pädagogik, 10*, 300–320.

Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., et al. ... Higgins, J. P. T. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *British Medical Journal, 343*, d4002. https://doi.org/10.1136/bmj.d4002.

Tanner-Smith, E. E., Tipton, E., & Polanin, J. R. (2016). Handling complex meta-analytic data structures using robust variance estimates: a tutorial in R. *Journal of Developmental and Life-Course Criminology, 2*, 85–112. https://doi.org/10.1007/s40865-016-0026-5.

Tipton, E. (2015). Small sample adjustments for robust variance estimation with meta-regression. *Psychological Methods, 20*, 375–393. https://doi.org/10.1037/met0000011.

Turner, J. E., & Schallert, D. L. (2001). Expectancy-value relationships of shame reactions and shame resiliency. *Journal of Educational Psychology, 93*, 320–329. https://doi.org/10.1037/0022-0663.93.2.320.

Tze, V. M. C., Daniels, L. M., & Klassen, R. M. (2016). Evaluating the relationship between boredom and academic outcomes: A meta-analysis. *Educational Psychology Review, 28*, 119–144. https://doi.org/10.1007/s10648-015-9301-y.

Vevea, J. L., & Woods, C. M. (2005). Publication bias in research synthesis: Sensitivity analysis using a prior weight functions. *Psychological Methods, 10*, 428–443. https://doi.org/10.1037/1082-989X.10.4.428.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48.

Weiner, B. (2007). Examining emotional diversity on the classroom: An attribution theorist considers the moral emotions. In P. A. Schutz, & R. Pekrun (Eds.). *Emotions in education* (pp. 73–88). San Diego, CA: Academic Press.

Wigfield, A., Rosenzweig, E. Q., & Eccles, J. S. (2017). Achievement values: Interactions, interventions, and future directions. In A. J. Elliot, C. S. Dweck, & D. S. Yeager (Eds.). *Handbook of competence and motivation* (pp. 116–134). (2nd ed.). New York, NY: Guilford.

Wine, J. D. (1971). Test anxiety and the direction of attention. *Psychological Bulletin, 76*, 92–104. https://doi.org/10.1037/h0031332.

Zeidner, M. (1998). *Test anxiety. The state of the art.* New York, NY: Plenum.

Zeidner, M. (2014). Anxiety in education. In R. Pekrun, & L. Linnenbrink-Garcia (Eds.). *International handbook of emotions in education* (pp. 265–288). New York, NY: Taylor & Francis.

Zeidner, M., & Matthews, G. (2005). Evaluative anxiety. In A. J. Elliot, & C. S. Dweck (Eds.). *Handbook of competence and motivation* (pp. 141–166). New York, NY: Guilford Press.