Educational Psychology Review (2021) 33:357–385
https://doi.org/10.1007/s10648-020-09545-x

META-ANALYSIS

Meta-Analysis of Emotional Designs in Multimedia Learning: A Replication and Extension Study

Rachel M. Wong[1] · Olusola O. Adesope[1]

Published online: 27 July 2020

Abstract
Emotions can both facilitate and hinder learning. Emotional design features such as colors and shapes can be embedded in multimedia learning environments to manipulate learners' affects and learning outcomes. However, some studies suggest that emotional designs promote learning, while others show that they hinder it. Although Brom et al. (Educational Research Review 25:100–119, Brom et al. 2018) published a meta-analysis on the use of emotional designs in multimedia learning, an updated search showed that more studies were published recently. Thus, the present meta-analysis is a replication and extension of Brom et al.'s (Educational Research Review 25:100–119, Brom et al. 2018) meta-analysis. A total of 28 articles yielded the following independent effect sizes for each outcome examined: retention (k = 28), transfer (k = 38), comprehension (k = 16), mental effort (k = 28), perceived difficulty (k = 19), change in positive affect (k = 25), intrinsic motivation (k = 28), and liking/enjoyment (k = 19). Results showed that including emotional designs enhanced learning outcomes (retention: $g+$ = 0.35; transfer: $g+$ = 0.27; comprehension: $g+$ = 0.29), change in positive affect ($g+$ = 0.09), intrinsic motivation ($g+$ = 0.15), mental effort ($g+$ = 0.11), liking/enjoyment ($g+$ = 0.10), and reduced perceived difficulty ($g+$ = − 0.21). Moderator analyses were conducted for retention, mental effort, intrinsic motivation, and positive affect, and findings showed that mean effect sizes were moderated by participant characteristics as well as methodological and contextual features of the studies. We discuss these findings as well as their theoretical and practical implications.

Keywords Emotional design · Meta-analysis · Anthropomorphic graphics · Colors

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10648-020-09545-x) contains supplementary material, which is available to authorized users.

✉ Olusola O. Adesope
olusola.adesope@wsu.edu

[1] Department of Kinesiology and Educational Psychology, College of Education, Washington State University, Pullman, WA 99164-2114, USA

Multimedia learning is defined as learning that occurs from both pictures and words (Mayer 2014). The majority of research on multimedia learning has focused primarily on the cognitive processes used to select, organize, and integrate information from a given source (Mayer 2014), and the impact of sensory, working, and long-term memory on these processes. In recent years, however, research examining the influence of learners' affect and motivation on learning in multimedia learning environments has proliferated. This has led to suggestions that affective factors are equally important to cognitive factors for enhancing learning. For example, previous findings indicate that positive emotions may influence leaners' cognitive processes, motivation, and academic achievement (Isen et al. 1987; Linnenbrink 2006; Mega et al. 2014), and that emotions evoked by certain colors and shapes may improve learning performance (Plass et al. 2014).

One area of research in which researchers have examined the impact of learners' emotions in multimedia learning environments is through the use of emotional designs. Emotional designs are design features that seek to influence learners' emotions to promote learning (Plass and Kaplan 2016) and can include how information is presented (e.g., the types of colors or shapes used) or the interactions in learning environments (e.g., the presence of a pedagogical agent to provide support; Plass and Kaplan 2016). Research on the use of emotional designs in multimedia learning pioneered by Um et al. (2012) raised attention on how key elements of learning materials that are aesthetically pleasing can increase learners' positive emotions, thus motivating students and facilitating learning. Findings from this study were corroborated by results from a recent meta-analysis conducted by Brom et al. (2018) that examined the effects of emotional designs such as anthropomorphic graphics and pleasant colors on learning performance and other cognitive and affective outcomes. The present study is a replication and an extension of the previous meta-analysis to update Brom et al.'s (2018) meta-analysis with more recent studies.

Emotional design research is built on well-established evidence on the relationship between emotions and learning achievement. Control-value theory (Pekrun 2006) posits that academic emotions related to achievement and academic events stem from learners' sense of control and autonomy over their academic experiences and the value they attribute to them. Specifically, the control-value theory describes emotions using two dimensions: valence, related to positive and negative states, and activation, related to activating and deactivating states (Pekrun 2006). This suggests that positive and negative emotions lead to different learning behaviors, attitudes, and outcomes. For example, Pekrun (2006) found that positive activating emotions such as happiness tend to enhance intrinsic and extrinsic motivation, while negative emotions can hinder learning outcomes. In addition, Brand et al. (2007) found that learners experiencing negative emotions often require more attempts for mastery performance and are more likely to perform poorly on transfer tasks than learners experiencing positive emotions.

The Cognitive Affective Theory of Learning with Media (CATLM; Moreno 2006) further elucidates the relationship between emotions and cognitive processes required for learning. The CATLM extends the Cognitive Theory of Multimedia Learning (CTML; Mayer 2014) to include motivation and affective factors. It is based on the following cognitive and affective assumptions:

a) there are separate channels for processing visual and auditory information (Paivio 1986; Baddeley 1992);
b) there is a limited amount of information that can be processed in each channel at one time (Baddeley 1992);

c) meaningful learning occurs when learners actively engage in selecting, organizing, and integrating incoming information with prior knowledge to form coherent mental representations (Mayer 2014; Wittrock 1989);
d) learning is mediated by motivational factors that affect learners' levels of cognitive engagement (Pintrich 2003);
e) learning is mediated by metacognitive strategies that regulate cognitive and affective processes (Moreno and Mayer 2007); and
f) individual differences affect the efficiency of learning with methods and media (Park et al. 2014).

According to the CATLM, learning takes place via three cognitive processes: selection, organization, and integration (Mayer 2014). In these processes, learners' motivation, affect, and metacognitive skills regulate any of the cognitive processes at any point (Moreno and Mayer 2007). Learning is also determined by three demands on cognitive processing: extraneous, essential, and generative processing (Mayer 2014; Moreno and Mayer 2007). When multimedia learning environments include extraneous material that is not relevant to the instructional objective, learners must engage in extraneous processing. This leaves fewer cognitive resources available for essential processing, which involves the cognitive demand of selecting information to be represented in working memory. It is influenced by the complexity of the task, as well as learners' levels of prior knowledge. Generative processing occurs when learners organize new information into coherent mental representations and integrate newly acquired mental representations with their prior knowledge.

In the context of this framework, emotional designs may either enhance the learning process by promoting generative processing (*emotion-as-facilitator-of-learning* hypothesis) or interfere with the process by increasing extraneous processing (*emotion-as-suppressor-of-learning* hypothesis; Um et al. 2012) due to increased demands on working memory to process the design features. The CATLM lays the foundation for examining how visually appealing elements regulate cognitive learning processes via learners' motivation and affect. Specifically, it provides strong theoretical support for the emotion-as-facilitator-of-learning hypothesis (Um et al. 2012), positing that positive emotions enhance learning processes by increasing motivation, which in turn improves generative processing, and thus learning performance (Mega et al. 2014; Pekrun 2006; Um et al. 2012).

The CATLM is further supported by Plass and Kalyuga (2019) who outlined four ways in which emotions relate to cognitive load theory: (1) emotions as a source of extraneous cognitive load; (2) emotions as a factor affecting memory; (3) emotions as intrinsic cognitive load; and (4) emotions affecting motivation to increase cognitive effort. Although a detailed discussion of Plass and Kalyuga's (2019) study is beyond the scope of this paper, it is clear that there are distinct overlaps between the four ways emotions relate to cognitive load theory and the CATLM. More specifically, the researchers appear to agree that emotions can both hinder and promote learning via learners' cognitive processing of information (Plass and Kalyuga 2019).

Emotional Designs in Multimedia Learning

Given the broad range of what constitutes emotional design, this meta-analysis, like Brom et al. (2018), focuses specifically on features that are used for information representation such as colors, shapes, and anthropomorphisms that are embedded within multimedia learning environments to induce positive emotions in learners (Park et al. 2015; Plass and Kaplan 2016; Um et al. 2012).

Research examining preferences for color revealed that most individuals prefer colors that are warm, highly saturated, and bright (Gorn et al. 1997; Palmer and Schloss 2010; Wolfson and Case 2000). Specifically, colors such as bright yellow and orange are more likely to elicit positive emotions in learners compared to colors such as black, brown, or gray (Palmer and Schloss 2010). Furthermore, studies that have examined colors as emotional design features (e.g., pleasant colors vs. black-and-white colors) provide some support for the continual use of aesthetically pleasing colors to enhance learning (e.g., Dong 2007; Plass et al. 2014; Um et al. 2012). While colors can also be used as signals or cues to direct learners' attention to key details in learning materials (Tabbers et al. 2004; Jamet et al. 2008), they should not be interpreted as emotional design features as they serve a different primary function. Specifically, colors used as signals or cues seek to *direct* attention and enhance cognitive processing with less regard for learners' emotions, while colors used as emotional design features primarily target learners' emotions, especially positive emotions (Plass and Kaplan 2016).

In addition to the use of colors as emotional design features, studies also show that round, human-like shapes are more likely to induce positive emotions than shapes with sharp edges (Berry and McArthur 1985; Plass et al. 2014). The addition of human-like features (e.g., eyes, mouths, and eyebrows) and shapes to nonhuman objects is termed anthropomorphism (Mayer and Estrella 2014; Waytz et al. 2010). Specifically, adding anthropomorphic graphics in learning materials has been found to increase learning performance (Brom et al. 2016; Schneider et al. 2018a, 2018b).

The incorporation of emotional designs in learning materials has focused largely on the use of pleasant colors and anthropomorphic graphics to invoke positive emotions, with few studies examining the impact of emotional designs on negative emotions (see Kumar 2016; Kumar et al. 2019). Despite the evidence suggesting that emotional designs are beneficial for learning, it is also important that these features are used to represent information relevant to the learning objective, rather than extraneous, seductive details that are interesting but not related to this objective (Harp and Mayer 1998). Although seductive details may increase interest in the learning materials, they are also likely to reduce learning outcomes and increase the extraneous cognitive load experienced by learners (Harp and Mayer 1998).

Purpose of the Present Meta-analysis

The aim of the present meta-analysis is to conduct a replication and extension of Brom et al.'s (2018) meta-analysis of anthropomorphism and pleasant colors in multimedia learning. Brom et al.'s (2018) meta-analysis is a classic study that provides a landscape of research in emotional design in multimedia learning environment. The researchers reviewed 33 independent articles (23 published and 10 unpublished) comparing the effects of some form of emotional design with a control condition without emotional designs. In general, they found that the use of emotional designs resulted in increases in retention, transfer, and comprehension of about $d = 0.39$, 0.32, and 0.33 standard deviations respectively. Emotional designs were also found to positively influence intrinsic motivation, liking/enjoyment, and positive affect, albeit producing smaller effects compared to learning outcomes. In general, the researchers concluded that anthropomorphic graphics and colors were useful designs to incorporate into a learning environment.

Although Brom et al.'s (2018) meta-analysis was a recent publication, a replication and extension of their study is important as a scientific approach to examine the robustness and provide validation of the findings in the existing meta-analysis (Lakens et al. 2016). In addition, since the publication of the previous meta-analysis, 8 new articles yielding 13 independent studies were eligible for inclusion in the present meta-analysis. Hence, the present meta-analysis seeks to answer the following research questions:

1. What are the effects of using emotional designs compared to neutral designs in multimedia learning?
2. How are emotional design effects moderated by presentation features?
3. How do emotional design effects vary with participant characteristics?
4. How are emotional design effects moderated by contextual features of studies?

Study Variables That Influence Learning Through Emotional Design

Type of Emotional Design Plass et al. (2014) observed that specific features of emotional designs have differential impacts on learning and affective outcomes. The use of anthropomorphic graphics and warm colors together or separately were found to increase comprehension. However, they were most effective for increasing transfer performance when they were used with neutral colors. Round, face-like shapes alone were found to induce positive emotions in learners, significantly improving comprehension. This suggests that using emotional design features with a combination of round, face-like shapes and colors, or round face-like shapes alone is more effective than using warm colors alone. Given this discrepancy in existing research, we examined type of emotional design as a potential moderator. We divided types of emotional design into four groups: (a) anthropomorphic graphics (e.g., Brom et al. 2016; Haaranen et al. 2015; Schneider et al. 2018a, 2018b), (b) colors (e.g., Dong 2007; Heidig et al. 2015), (c) anthropomorphic graphics and colors (e.g., Park et al. 2015; Plass et al. 2014; Um et al. 2012), and (d) mixed (e.g., Gong et al. 2017, exp 2; Plass et al. 2014, exp 2; Uzun and Yildirim 2018).

Several studies made additional manipulations such as the sharpness/roundness of edges (Münchow et al. 2017), facial expressions of the human character (Uzun and Yildirim 2018), and the presence of arrows in the neutral condition to depict direction (Mayer and Estrella 2014). However, since these were minor alterations to the learning materials and these additions varied across studies, like Brom et al. (2018), we included these studies in the meta-analysis but only coded for the presence of one of the four primary emotional design type.

Emotional Induction The effectiveness of emotional design features in instructional materials is also determined by learners' affective states before learning occurs. Münchow et al. (2017) observed that learners' pre-experimental affective states were as important for predicting learning outcomes as the emotions induced through the learning intervention. Learners with high positive affect prior to learning performed significantly better on transfer problems when they also viewed positive emotional designs in the learning materials. This effect can be explained by the fact that positive emotions promote generative processing (Pekrun 2006). Therefore, emotions experienced prior to learning have a facilitating effect during the learning process, enhancing learning gains. Hence, learners' initial affective states prior to learning with emotional designs are likely to influence the extent of the effects of emotional designs on learning outcomes. To better understand the effects of initial affective states, we determined the presence of emotional induction as a potential moderator.

Pacing The pacing of the learning instruction and material is classified as either learner-paced or system-paced, depending on whether learners had control over the speed of the presented material. Specifically, for learner-paced instruction, learners have the opportunity to pause and continue the presentation based on their cognitive and learning needs, while system-paced instruction are timed and progress automatically. Some research suggests that giving learners control over the pacing of the material can be beneficial for transfer of knowledge (Hasler et al. 2007; Tabbers and de Koeijer 2010). Yet at the same time, this finding has not always been supported or deemed universally advantageous due to several factors including individual differences in learners (Höffler and Schwartz 2011) and the design of the learning environment (Rop et al. 2018). In terms of affective outcomes, a recent study by Shangguan, Wang, Gong, Guo, and Xu (Shangguan et al. 2019) found that emotional designs had a greater impact on participants' positive emotions when learners were also given control over the pace of their learning material. To provide a more conclusive understanding of the learning benefits of pacing in multimedia learning, this meta-analysis examined the effects of pacing on emotional, motivation, and learning outcomes.

Nature of Presentation In this meta-analysis, the nature of presentation is defined as animated or static. Previous meta-analyses comparing animated with static graphics show that animation enhances learning more than static graphics ($g = 0.23$, Berney and Bétrancourt 2016; $d = 0.37$, Höffler and Leutner 2007). This may be due to two factors. First, the directing function of animations likely cues learners' attention to pertinent information in the learning materials (Lowe & Schnotz 2014). Animation may emphasize key ideas that a learner may otherwise struggle to identify on their own. Second, since animated materials aid in the visualization of learning materials (Höffler and Leutner 2007), studying such material may reduce learners' cognitive load. To elucidate how the nature of presentation affects learning, we examined the use of animation as a moderator. In some animation studies, participants viewed learning materials in a series of frames that evoked a sense of continuous change in the imagery (Lowe & Schnotz 2014), while participants in static studies viewed learning materials in a single static frame.

Educational Level and Language/Culture Studies on the use of emotional designs in multimedia learning have been conducted with participants of diverse age groups, ranging from elementary-aged students (Ng and Chiu 2017) to postsecondary-aged students (Dong 2007; Um et al. 2012), and diverse cultural backgrounds including German (Münchow 2017; Park et al. 2015; Plass et al. 2014), Chinese (Gong et al. 2017), and American (Dong 2007; Miller 2011; Um et al. 2012). We coded participants' educational level from each study based on reported age/grade level: elementary, middle, high, or postsecondary. We coded for language/culture based on where studies were conducted. When the sample size for specific cultures was less than three, we grouped those studies under "other."

Study Setting Emotional design studies are often conducted in either controlled laboratory or classroom settings. While most studies in this meta-analysis were laboratory-based (k = 38), seven of the coded studies were conducted in a classroom setting. We defined a classroom study as one that utilized materials from the participants' regular in-class curriculum (e.g., Kumar et al. 2016, 2019). Studies that utilized generic learning materials that did not align with students' curriculum were coded as laboratory studies. Since students in classroom studies received material that aligned with their course objectives, they may have higher levels of motivation to learn as compared to students in controlled laboratory studies that have little bearing on their overall academic performance. Hence, we aimed to establish the degree to which setting moderated the effects of emotional designs.

Learning Domain We coded learning domain as a categorical variable with the following five categories: (a) life science, (b) physical science, (c) computer science, or (d) others. The term life science was used to capture any biology-related topics such as immunization, functional neuroanatomy, and blood cells, and physical science was used to capture any topics related to weather or meteorology.

Instructional Time We coded the time of treatment implementation in terms of minutes and then divided that into categories based on treatment time: short (less than 15 min), moderate (15 to 30 min), or long (greater than 30 min).

Publication Source The publication source of each study examined served as an indicator of study quality. Studies were coded based on the source that they were obtained from: journal, conference paper, book chapter, or dissertation.

These nine study variables identified in the present meta-analysis were selected based on our review of the literature and are also identical to the variables identified in Brom et al.'s (2018) meta-analysis. It is not surprising that these variables may contribute to inconclusive results on the effectiveness of emotional designs in multimedia learning, since research on emotional designs in multimedia learning have been conducted with different populations, comparison treatments, and outcome measures. Thus, this meta-analysis is warranted to comprehensively examine the landscape of research in this area and reconcile mixed findings.

Method

Selection Criteria

For this meta-analysis, a study was deemed eligible for inclusion if it:

1. contrasted the effects of a positive emotional design condition with a neutral emotional design condition in which emotional designs (e.g., color and/or anthropomorphic graphics) are embedded in the learning intervention;
2. reported measurable learning outcomes such as retention, transfer, or comprehension;
3. reported measurable affective outcomes such as motivation, satisfaction, and change in affect;
4. reported measurable perceptions of cognitive load/mental effort

5. reported sufficient data for effect size extractions;
6. randomly assigned participants to groups, or reported pre-test or other prior data to control for pre-existing differences between groups; and
7. was publicly available through online databases, journals, or library archives.

Studies were excluded if they did not report enough descriptive statistics for effect size calculations. When there were multiple reports of the same study (e.g., dissertation, journal article, conference proceedings), the journal publications were used; however, other versions (e.g., dissertation or conference proceedings) were used to support and extend coding features (see Um 2008; Um et al. 2012). For example, Brom et al.'s (2018) meta-analysis reported dissertations by both Kumar (2016) and Münchow (2017); however, at the time of our search, these dissertations had been published as journal articles. Therefore, the journal publications were included in the present meta-analysis whereas the dissertations provided additional information for data extraction.

Location and Selection of Studies

We used the query *emotion** AND *multimedia* AND *learn** to conduct a comprehensive and systematic search on the following electronic databases: ERIC, ProQuest, PsycINFO, PsycARTICLES, and Web of Science. Although the search query was different from what was used in Brom et al.'s (2018) meta-analysis, the resulting articles were the same. The database search completed in February 2020 returned a total of 878 articles.

There were two screening phases. In the first phase, we read the titles and abstracts of all 878 articles to determine possible inclusion in the meta-analysis. A total of 49 articles met all inclusion criteria and were retained in the second phase. Full-text copies of the 49 articles were obtained. In the second screening phase, the 49 articles were reviewed by applying the selection criteria listed above. Out of the 49 articles, a total of 19 articles met all inclusion criteria. A secondary examination of the reference lists included articles, Brom et al.'s meta-analysis, and feedback provided by reviewers returned an additional 9 articles that fit the inclusion criteria. Figure 1 provides a visual depiction of the inclusion/exclusion criteria.

Data from the 28 studies were extracted and entered into a master coding form. The coded variables were organized by the following categories: (a) study identification, (b) research questions, (c) sample information, (d) treatment and control condition, (e) research design, (f) dependent variables, and (g) results. When coded variables were not explicitly reported, the researcher attempted to contact the corresponding author of the publication for clarification. To assess reliability, one of the authors randomly selected and coded 20% of all studies included in this meta-analysis. The mean inter-coder agreement was 98.3%, indicating high consistency between the two coders.

Extraction and Calculation of Effect Sizes

The standardized mean difference, Cohen's *d*, served as the measure of effect size; the difference between the experimental (emotional design material) and the control (neutral design material) group scores divided by the pooled standard deviations of the two groups. Cohen's *d* was computed for each independent study. To obtain Cohen's *d* for positive affect, we calculated the change in pre and post affect scores (post–pre), as separate pre and post affect scores were more commonly reported than change score. Notably, this was different

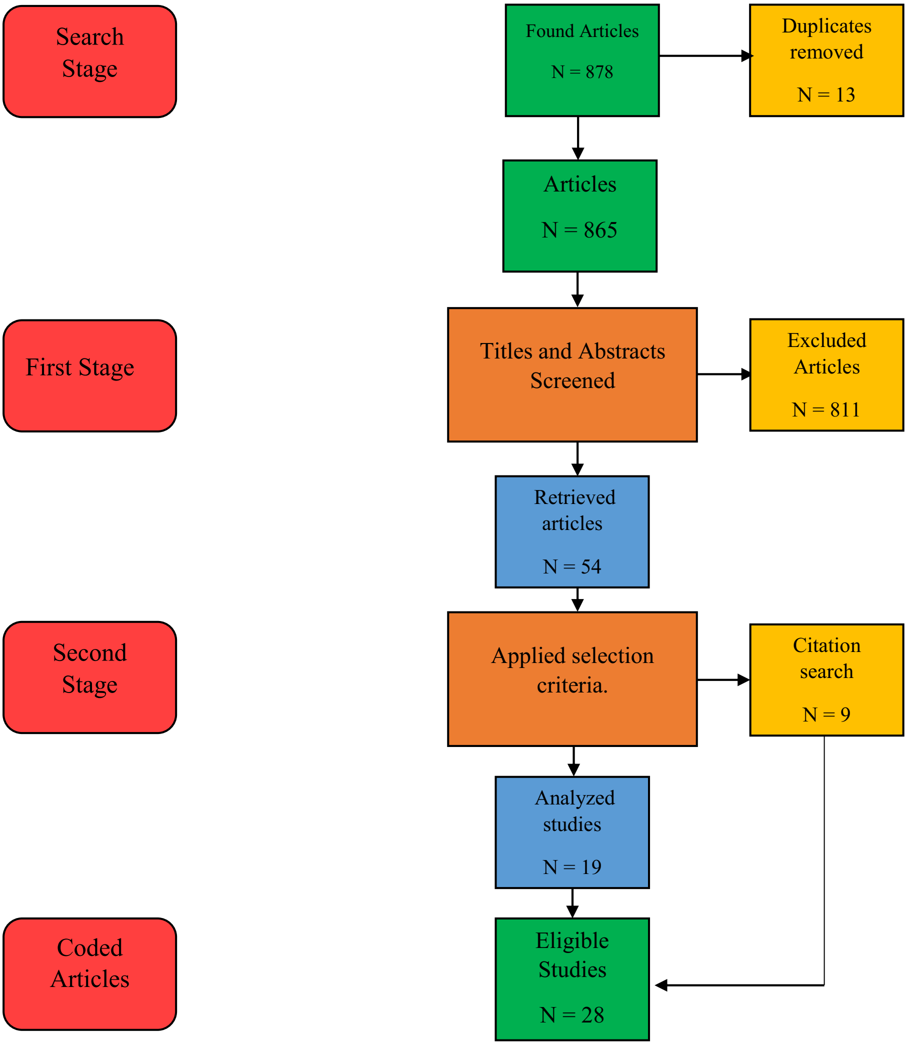

Fig. 1 Visual depiction of the selection of studies

than what was done in Brom et al.'s (2018) meta-analysis. As some articles reported multiple studies, the 28 primary research articles/dissertations yielded the following independent effect sizes for each outcome examined: retention ($k = 28$), transfer ($k = 38$), comprehension ($k = 16$), mental effort ($k = 28$), perceived difficulty ($k = 19$), change in positive affect ($k = 25$), intrinsic motivation ($k = 28$), and liking/enjoyment ($k = 19$).

Due to the inherent bias in standardized mean difference effect sizes with small sample studies, Hedges' g was computed (Hedges 1981) from Cohen's d. Hedges' g was reported in this meta-analysis as the unbiased estimate of the standardized mean difference effect. When means or standard deviations were not reported, formulas for calculating d from t or F statistics were used. A positive g indicates an advantage of using emotional designs over neutral designs, and a negative g indicates an advantage of using neutral designs over emotional designs.

Multiple experiments within a single journal article were treated as independent studies (e.g., Gong et al. 2017; Mayer and Estrella 2014; Schneider et al. 2018a, 2018b). Like Brom et al. (2018), when studies used 2 × 2 designs, we were only interested in the primary factor in which emotional designs were manipulated. Specifically, this meant that each level of the secondary factor (e.g., mood induction: positive or neutral; or behavioral positive design: learner-control or system-control) was treated as an independent sample. For example, in a 2 × 2 factorial design, Plass et al. (2014) examined the effects of emotional designs (presence of emotional designs vs. absence of emotional designs) and prior mood induction (positive vs. neutral) on learning, cognitive, and affective outcomes. Treating the two levels of prior mood induction (positive and neutral) as independent samples resulted in two sets of comparisons (e.g., PEPD vs. PEND and NEND vs. NEPD) and thus, two sets of effect sizes.

With regard to data treatment for studies that compared multiple experimental conditions against a single control group, a weighted mean and pooled standard deviation was calculated for relevant experimental conditions in each study. Doing so helps maintain independence in the meta-analysis and ensures that the control condition is only entered once (Borenstein, Hedges, Higgins, & Rotherstein, 2013). This was done specifically for studies such as Heidig et al. (2015) that had eight experimental conditions and one control group, and Gong et al. (2017; exp 2) and Plass et al. (2014; exp 2) that both had three experimental conditions and one control group.

Data Analysis

Data were analyzed using Comprehensive Meta-Analysis version 3.0 (Borenstein et al. 2013) and IBM SPSS version 24 for Windows. Weighted mean effect sizes were also calculated to assign more weight to over-represented studies (*g*+). Prior to data analysis, all effect sizes were examined for outliers and results indicated the absence of outliers. All effect sizes were included for further computation. The *Q* statistic generated by the Comprehensive Meta-Analysis software was used to examine homogeneity of variance of the effect sizes. In addition to the *Q* statistics, the I^2 statistic was also used to interpret the result of the homogeneity test. The I^2 statistic represents variability percentage attributed to heterogeneity and is reportedly a more comprehensive measure (Higgins and Thompson 2002). Therefore, both *Q* and I^2 statistics, generated by the Comprehensive Meta-Analysis software, were used to provide robust information on the amount of homogeneity of variance across studies. The presence of heterogeneity was determined with a significant *Q* statistic (e.g., $p < .05$) and is indicative of moderators that may influence the effect sizes from the studies. In addition to the *Q* statistic, the I^2 statistic was used to determine heterogeneity based on the following recommended criteria (Higgins and Thompson 2002): percentages around 25%, 50%, and 75% represented low, medium, and high heterogeneity respectively. Finally, publication bias was examined via Begg and Mazumdar's rank correlation test and Egger's linear regression test.

Results

A total of 28 primary studies yielded independent effect sizes for the following: retention ($k = 28$), transfer ($k = 38$), comprehension ($k = 16$), mental effort ($k = 28$), perceived difficulty ($k = 19$), intrinsic motivation ($k = 28$), liking/enjoyment ($k = 19$), and change in positive affect ($k = 25$).

These effect sizes were analyzed on different features and outcomes. Separate meta-analyses were conducted for each outcome.

Emotional Design Literature

All tables of results include the following data: (a) number of participants (*N*) in each category, the number of findings (*k*), the weighted mean effect size (*g*+), standard error (*SE*), the 95% lower and upper confidence intervals (95% CI), the results of a test of homogeneity (*Q*) along with degrees of freedom (*df*), and percentage of variability attributed to heterogeneity (I^2). Table 1 shows the overall analysis of the weighted means of the effect sizes for learning, cognitive, and affective outcomes under a fixed-effects model. Positive weighted mean effect size indicates enhanced outcomes for the emotional design treatments compared to the neutral design treatments.

Despite the small to moderate statistically detectable effect sizes for comprehension (*g*+ = 0.29), perceived difficulty (*g*+ = − 0.21), and liking/enjoyment (*g*+ = 0.10), there was no significant variability between studies for these outcomes, as indicated by the homogenous distribution. Specifically, the tests of homogeneity for the above outcomes are as follows: comprehension, $Q(15) = 17.32$, $p = .29$, $I^2 = 13.88$; perceived difficulty, $Q(18) = 21.82$, $p = .24$, $I^2 = 17.51$; and liking/enjoyment, $Q(18) = 21.21$, $p = .27$, $I^2 = 15.12$.

On the other hand, results showed small to moderate statistically significant effects of learning with emotional designs for retention (*g*+ = 0.35), transfer (*g*+ = 0.27), mental effort, (*g*+ = 0.11), intrinsic motivation (*g*+ = 0.16), and liking/enjoyment (*g*+ = 0.16). However, the distributions for these outcomes were heterogeneous–retention: $Q(27) = 120.95$, $p < .001$, $I^2 = 77.68$; transfer: $Q(37) = 78.12$, $p < .001$, $I^2 = 52.64$; mental effort: $Q(27) = 95.29$, $p < .001$, $I^2 = 71.67$; intrinsic motivation: $Q(27) = 82.80$, $p < .001$, $I^2 = 67.39$; and change in positive affect: $Q(24) = 53.09$, $p < .001$, $I^2 = 54.79$.

The total variability that could be attributed to true heterogeneity for retention, transfer, mental effort, intrinsic motivation, and change in positive affect was 77.8%, 52.6%, 71.7%, and 64.1% respectively. This indicates that over 60% of the variance for four out of five outcome variables was due to between-studies variance and could be explained by study-level covariates, and less than 40% of the variance was within-study variance. Since these variables had moderate to high heterogeneous distribution of these outcomes (Higgins and Thompson

Table 1 Weighted mean effect sizes for learning, cognitive, and affective outcomes

	Effect size					Test of heterogeneity			
Variable	*N*	*k*	*g*+	SE	95% CI	*Q*	df	*p*	I^2 (%)
Retention	2434	28	0.35	0.04	[0.26, 0.43]	120.95	27	< .001	77.68
Transfer	3016	38	0.27	0.04	[0.19, 0.35]	78.12	37	< .001	52.64
Comprehension	1903	16	0.29	0.05	[0.20, 0.39]	17.42	15	.29	13.88
Mental effort	1699	28	0.11	0.05	[0.01, 0.21]	95.29	27	< .001	71.67
Perceived difficulty	1296	19	− 0.21	0.06	[− 0.32, − 0.10]	21.82	18	.24	17.51
Intrinsic motivation	2411	28	0.15	0.04	[0.06, 0.23]	82.80	27	< .001	67.39
Liking/enjoyment	1409	19	0.10	0.05	[− 0.003, 0.21]	21.21	18	.27	15.12
Change in positive affect	2157	25	0.09	0.05	[0.01, 0.18]	53.09	24	< .001	54.79

*$p < .05$

2002), subsequent moderator analyses were performed to identify potential moderators for the retention, transfer, mental effort, intrinsic motivation, and change in positive affect.

Moderator Analyses

Notably, many of the studies included in our meta-analysis were conducted by few researchers who used similar learning materials and design conditions to examine the effect of emotional designs on learning, cognitive, and affective outcomes. As such, in this meta-analysis, we assumed that the differences in observed effects are due to sampling error (Borenstein et al. 2009). Unlike Brom et al. (2018) who utilized a random-effects model, a fixed-effects model was used for subsequent moderator analyses in this present meta-analysis. Multiple comparisons were made for each moderator analysis within each outcome variable. To correct for multiple comparisons and to prevent type I error of obtaining false positives, we adjusted the *p* value of each of the moderators for each outcome variable. Similar to Brom et al. (2018), we applied the false-discovery rate (FDR) adjustment (Benjaminni and Hochberg 1995) as it was the most feasible approach. Although transfer had a low to moderate heterogeneous distribution, subsequent analysis revealed that transfer did not produce any significant moderators of interest. Therefore, the following paragraphs provide the results of the moderator analyses for retention, mental effort, intrinsic motivation, and positive affect.

Retention Table 2 summarizes the findings for moderator analysis within retention. Results showed that for presentation features, the only significant moderator was type of emotional design, $Q(3) = 10.64$, $p = .046$. Specifically, studies that utilized anthropomorphic shapes, anthropomorphic shapes + colors, and colors alone were associated with small to moderate effect sizes, $g+ = 0.55$, $g+ = 0.28$ and $g+ = 0.46$ respectively. Post hoc comparisons revealed that the use of anthropomorphic shapes alone differed significantly from studies that used mixed emotional design types; however, anthropomorphic shapes did not differ significantly from the use of colors alone. Notably, given the similar confidence intervals for anthropomorphic shapes and colors, one might infer that anthropomorphic graphics alone or colors alone may not be more beneficial than the other for learning. No other moderators were significant for presentation features.

For participant characteristics, grade was the only significant moderator, $Q(3) = 73.39$, $p < .001$. Post hoc comparisons revealed that studies conducted with elementary school aged students ($g+ = 2.15$) were associated with a much larger effect size that differed significantly from studies conducted with middle school students ($g+ = 0.28$), high school students ($g+ = 0.53$), and postsecondary students ($g+ = 0.20$). While no other significant comparisons were detected, it is necessary to point out that a majority of studies were conducted with postsecondary students and fewer students with k-12 students. This necessitates careful interpretation of the results, but more importantly highlights the need for more studies to be conducted with the k-12 population.

Finally, learning domain was found to be the only significant moderator for contextual features, $Q(2) = 45.10$, $p < .001$. Studies that were coded as others ($g+ = 1.20$) were associated with a significantly larger effect size compared to studies in the life ($g+ = 0.26$) and physical sciences ($g+ = 0.22$). The similar confidence intervals for studies in the life sciences and studies in the physical sciences indicate that the use of emotional designs in either domain was not more beneficial than the other.

Table 2 Weighted mean effect sizes within retention

Moderator	*N*	*k*	*g*+	SE	95% CI	$Q_{between}$	df	$p_{adjusted}$
Presentation features								
Type of emotional design								
Anthropomorphic shapes	544	6	0.55**	0.09	[0.37, 0.72]			
Anthropomorphic shapes w/ colors	1125	15	0.28**	0.06	[0.16, 0.40]			
Colors	477	4	0.46**	0.12	[0.22, 0.70]			
Mixed	288	3	0.10	0.13	[− 0.15, 0.35]			
Between levels (Q_B)						10.64	3	.046
Prior emotion induction								
Yes	188	4	0.04	0.15	[− 0.24, 0.33]			
No	2246	24	0.37**	0.05	[0.29, 0.46]			
Between levels (Q_B)						4.69	1	.08
Nature of presentation								
Animated	1372	17	0.41**	0.06	[0.30, 0.53]			
Static	1062	11	0.27**	0.06	[0.15, 0.39]			
Between levels (Q_B)						2.71	1	.12
Pacing of presentation								
Learner-paced	1636	15	0.32**	0.05	[0.21, 0.43]			
System-paced	721	11	0.40**	0.08	[0.25, 0.56]			
Not reported	77	2	0.27	0.23	[− 0.19, 0.72]			
Between levels (Q_B)						0.85	2	.73
Participant characteristics								
Grade								
Elementary school	122	2	2.15**	0.23	[1.71, 2.60]			
Middle school	515	7	0.28*	0.09	[0.11, 0.46]			
High school	376	3	0.53**	0.11	[0.31, 0.74]			
Postsecondary	1421	16	0.20**	0.06	[0.08, 0.31]			
Between levels (Q_B)						73.39	3	< .001
Language/culture								
American	254	5	0.52*	0.13	[0.27, 0.77]			
Chinese	747	11	0.36*	0.08	[0.21, 0.51]			
Czech	248	3	0.08	0.13	[− 0.18, 0.33]			
German	1104	8	0.33*	0.07	[0.19, 0.46]			
Other	81	1	0.71	0.24	[0.24, 1.18]			
Between levels (Q_B)						8.76	4	.11
Contextual features								
Study settings								
Classroom	147	2	0.63**	0.17	[0.29, 0.97]			
Laboratory	2287	26	0.33**	0.05	[0.24, 0.41]			
Between levels (Q_B)						2.97	1	.12
Instructional time								
Short	771	12	0.45**	0.07	[0.30, 0.59]			
Moderate	541	4	0.26*	0.09	[0.08, 0.43]			
Long	381	2	0.58**	0.15	[0.28, 0.88]			
Not reported	741	10	0.24*	0.08	[0.09, 0.39]			
Between levels (Q_B)						7.13	3	.11
Learning domain								
Life science	1108	11	0.26**	0.06	[0.14, 0.39]			
Physical science	1057	12	0.22*	0.07	[0.08, 0.36]			
Other	269	5	1.20**	0.13	[0.93, 1.46]			
Between levels (Q_B)						45.10	2	< .001
Publication								
Conference	67	1	0.49*	0.25	[0.01, 0.98]			
Dissertation	77	2	0.27	0.23	[− 0.19, 0.72]			
Journal	2290	25	0.34**	0.05	[0.25, 0.43]			
Between levels (Q_B)						0.49	2	.78

Note: *$p < .05$; **$p < .001$

Mental Effort Table 3 presents results for the moderator analysis within mental effort. With regard to presentation features, three out of four moderators of interest were significant: type of emotional design, $Q(3) = 22.19$, $p < .001$; prior emotion induction, $Q(1) = 5.42$, $p = .049$; and nature of presentation, $Q(1) = 9.60$, $p = .007$. The use of anthropomorphic shapes was associated with a large significant positive effect size ($g+ = 1.02$) that differed significantly from the use of anthropomorphic shapes + colors ($g+ = 0.11$), colors alone ($g+ = -0.36$) and mixed emotional designs ($g+ = 0.12$). Notably, the use of colors alone was associated with a negative effect size, indicative that colors were not as beneficial as the use of the other emotional designs in helping increase participants' mental effort. Given the uneven distribution of studies across the four types of emotional designs, we highly recommend caution in the interpretation of this finding. For prior emotion induction, studies that did not induce participants' emotions prior to the intervention were associated with a small effect size ($g+ = 0.20$). However, post hoc comparisons indicated that this did not differ significantly from studies that induced emotions prior to the intervention. Finally, for nature of presentation, studies that used static materials were association with a moderate effect size ($g+ = 0.41$) that were significantly different from the studies that used animated presentations.

For participant characteristics, grade level was the only significant moderator, $Q(1) = 10.06$, $p = .007$. Studies conducted with middle school participants were associated with a small effect size ($g+ = 0.38$) that was significantly different from studies conducted with postsecondary participants ($g+ = 0.02$). However, at the same time, we recognize that majority of the studies ($k = 22$) were conducted with postsecondary participants and few studies were conducted with middle school participants ($k = 6$). Therefore, it would be advisable to practice caution in the interpretation of these results. No other significant moderators were observed for mental effort.

Intrinsic Motivation Table 4 summarizes the results for the moderator analysis for intrinsic motivation. Under presentation features, results showed that three of four moderators were significant: type of emotional design, $Q(3) = 34.39$, $p < .001$; nature of presentation, $Q(1) = 17.53$, $p < .001$; and pacing, $Q(1) = 30.11$, $p < .001$. Specifically, for type of emotional design, post hoc comparisons showed that the use of anthropomorphic shapes was associated with a moderate effect size ($g+ = 0.64$) that differed significantly from the use of anthropomorphic shapes + colors ($g+ = 0.004$), colors alone ($g+ = 0.17$), and mixed emotional designs ($g+ = -0.13$). No other significant comparisons were observed. For nature of presentation, the results indicated that materials presented statically were associated with a moderate effect size ($g+ = 0.42$) that differed significantly from animated presentations ($g+ = 0.03$). Finally, learner-paced studies were associated with a small positive effect size ($g+ = 0.31$) that differed significantly from system-paced studies that had a small negative effect size ($g+ = -0.20$).

Both moderators for participant characteristics were significant: grade, $Q(3) = 21.45$, $p = .005$ and language/culture, $Q(4) = 26.83$, $p < .001$. Specifically, for grade, studies that were conducted with both middle school and high school participants were associated with small positive effect sizes ($g+ = 0.20$ and $g+ = 0.46$, respectively). Post hoc comparisons indicated that there were significant differences between studies conducted with high school participants compared to studies conducted with elementary school participants as well as postsecondary participants. However, as previously mentioned, due to the uneven distribution of studies across grade levels, it is recommended to interpret these results with caution. Results for language/culture showed that studies with American participants were associated with a moderate effect size ($g+ = 0.55$) while studies with German participants conducted were associated with a small effect size ($g+ = 0.27$). Post hoc comparisons revealed that both

Table 3 Weighted mean effect sizes within mental effort

Moderator	*N*	*k*	*g*+	SE	95% CI	$Q_{between}$	df	$p_{adjusted}$
Presentation features								
Type of emotional design								
Anthropomorphic shapes	81	2	1.02**	0.24	[0.55, 1.48]			
Anthropomorphic shapes w/ colors	1084	19	0.11	0.06	[− 0.01, 0.23]			
Colors	143	3	− 0.36*	0.17	[− 0.70, − 0.02]			
Mixed	391	4	0.12	0.11	[− 0.10, 0.34]			
Between levels (Q_B)						22.19	3	< .001
Prior emotion induction								
Yes	562	12	− 0.05	0.09	[− 0.22, 0.12]			
No	1137	16	0.20*	0.06	[0.08, 0.32]			
Between levels (Q_B)						5.42	1	.049
Nature of presentation								
Animated	1326	22	0.03	0.06	[− 0.08, 0.14]			
Static	373	6	0.41**	0.11	[0.20, 0.62]			
Between levels (Q_B)						9.60	1	.007
Pacing of presentation								
Learner-paced	922	15	0.09	0.07	[− 0.05, 0.22]			
System-paced	70	11	0.13	0.08	[− 0.02, 0.28]			
Not reported	77	2	0.21	0.22	[− 0.23, 0.65]			
Between levels (Q_B)						0.38	2	.90
Participant characteristics								
Grade								
Middle school	428	6	0.38**	0.10	[0.19, 0.58]			
Postsecondary	1271	22	0.02	0.06	[− 0.10, 0.13]			
Between levels (Q_B)						10.06	1	.007
Country								
American	406	9	− 0.04	0.10	[− 0.24, 0.16]			
Chinese	625	9	0.15	0.08	[− 0.02, 0.31]			
Czech	181	2	0.20	0.15	[− 0.10, 0.50]			
German	406	7	0.09	0.10	[− 0.11, 0.30]			
Other	81	1	0.54*	0.24	[0.08, 1.01]			
Between levels (Q_B)						6.02	4	.28
Contextual features								
Study settings								
Classroom	147		2	− 0.21	0.18	[− 0.55, 0.14]		
Laboratory	1552		26	0.14*	0.05	[0.04, 0.24]		
Between levels (Q_B)							3.52	1 .12
Instructional time								
Short	582	9	0.26	0.08	[0.09, 0.42]			
Moderate	376	7	− 0.07	0.11	[− 0.29, 0.14]			
Long	47	1	− 0.10	0.29	[− 0.67, 0.47]			
Not reported	694	11	0.10	0.08	[− 0.06, 0.26]			
Between levels (Q_B)						6.44	3	.15
Learning domain								
Life science	829	14	0.04	0.07	[− 0.10, 0.18]			
Physical science	723	11	0.20*	0.08	[0.05, 0.35]			
Other	147	3	0.06	0.18	[− 0.28, 0.41]			
Between levels (Q_B)						2.38	2	.38
Publication								
Conference	34	2	0.12	0.42	[− 0.70, 0.95]			
Dissertation	77	2	0.21	0.23	[− 0.23, 0.66]			
Journal	1588	24	0.11*	0.05	[0.01, 0.21]			
Between levels (Q_B)						0.21	2	.90

Note: $*p < .05$; $**p < .001$

Table 4 Weighted mean effect sizes within intrinsic motivation

Moderator	*N*	*k*	*g*+	SE	95% CI	$Q_{between}$	df	$p_{adjusted}$
Presentation features								
Type of emotional design								
Anthropomorphic shapes	477	5	0.64**	0.10	[0.45, 0.83]			
Anthropomorphic shapes w/ colors	1163	17	0.004	0.06	[− 0.11, 0.12]			
Colors	561	4	0.17	0.11	[− 0.04, 0.38]			
Mixed	210	2	− 0.13	0.16	[− 0.44, 0.18]			
Between levels (Q_B)						34.39	3	< .001
Prior emotion induction								
Yes	451	8	0.09	0.09	[− 0.09, 0.27]			
No	1960	20	0.16**	0.05	[0.07, 0.26]			
Between levels (Q_B)						0.50	1	.48
Nature of presentation								
Animated	1691	21	0.03	0.05	[− 0.08, 0.13]			
Static	720	7	0.42**	0.08	[0.27, 0.57]			
Between levels (Q_B)						17.53	1	< .001
Pacing of presentation								
Learner-paced	1713	17	0.31**	0.05	[0.21, 0.42]			
System-paced	698	11	− 0.20*	0.08	[− 0.35, − 0.05]			
Between levels (Q_B)						30.11	1	< .001
Participant characteristics								
Grade								
Elementary school	122	2	− 0.20	0.18	[− 0.55, 0.15]			
Middle school	434	6	0.20*	0.10	[0.01, 0.39]			
High school	407	3	0.46**	0.11	[0.26, 0.67]			
Postsecondary	1448	17	0.10	0.06	[− 0.02, 0.21]			
Between levels (Q_B)						21.45	3	< .001
Language/culture								
American	184	3	0.55**	0.15	[0.26, 0.85]			
Chinese	687	10	− 0.16*	0.08	[− 0.30, − 0.004]			
German	1281	12	0.27**	0.06	[0.15, 0.40]			
Other	259	3	0.17	0.12	[− 0.07, 0.42]			
Between levels (Q_B)						26.83	4	< .001
Contextual features								
Study settings								
Classroom	227	3	0.35*	0.13	[0.08, 0.61]			
Laboratory	2184	25	0.12*	0.05	[0.03, 0.22]			
Between levels (Q_B)						2.44	1	.13
Instructional time								
Short	580	9	− 0.12	0.08	[− 0.28, 0.04]			
Moderate	883	9	0.40**	0.07	[0.26, 0.54]			
Long	593	4	0.07	0.10	[− 0.13, 0.27]			
Not reported	355	6	0.12	0.11	[− 0.10. 0.34]			
Between levels (Q_B)						23.61	3	< .001
Learning domain								
Life science	1082	12	0.31**	0.06	[0.19, 0.43]			
Physical science	269	11	− 0.05	0.07	[− 0.18, 0.09]			
Other	1060	5	0.16	0.12	[− 0.08, 0.40]			
Between levels (Q_B)						14.62	2	< .001

Note: *$p < .05$; **$p < .001$

American and German cultures differed significantly from studies conducted in Chinese ($g+ = -0.16$); however, no other significant comparisons were observed.

Two out of three moderators for contextual features were found to be significant as well: instructional time, $Q(3) = 23.61$, $p < .001$ and learning domain, $Q(2) = 14.62$, $p < .001$. For instructional time, studies coded as having moderate instructional time (between 15 and 30 min) were associated with a small effect size ($g+ = 0.40$). Post hoc comparisons indicate that studies with moderate instructional time differed significantly from studies that had short instructional time; however, no other significant comparisons were observed. For learning domain, both studies coded as life science and other were associated with moderate effect sizes ($g+ = 0.31$ and $g+ = 0.16$, respectively). Post hoc comparisons showed that while these two domains did not differ significantly from each other, they each differed significantly from studies that utilized learning materials in the physical science domain. Finally, publication type was not included as a moderator for intrinsic motivation as all studies that examined intrinsic motivation were published in journals.

Change in Positive Affect Table 5 summarizes the results for the moderator analysis for change in positive affect. For presentation features, the type of emotional design was the only significant moderator, $Q(3) = 12.03$, $p = .02$. Studies that used anthropomorphic shapes + colors were associated with a small positive effect size ($g+ = 0.19$), while studies that used colors alone were associated with a small negative effect size ($g+ = - 0.21$). Post hoc comparisons revealed significant differences between the use of anthropomorphic shapes + colors and colors alone; however, no other significant comparisons were observed. Language/culture was the only significant moderator for participant characteristics, $Q(4) = 16.01$, $p = .01$. Studies with Chinese participants were associated with a small effect size ($g+ = 0.35$); however, post hoc analysis revealed no significant comparisons between studies with Chinese participants and studies with other participants. Finally, instructional time was the only significant moderator for contextual features, $Q(3) = 26.52$, $p < .001$. Studies with short and moderate instructional times were associated with small positive effect sizes ($g+ = 0.31$ and $g+ = 0.22$, respectively); however, studies with long instructional times were associated with a moderate negative effect size ($g+ = -0.49$). The overlapping confidence intervals for short and moderate instructional times indicate that a moderate exposure to the learning material may not be more beneficial than a short exposure to the material. However, post hoc analysis revealed significant comparisons between long instructional times and both short and moderate instructional times.

Publication Bias

Studies with non-significant findings are either not published as frequently as studies with significant findings (Franco et al. 2014; Rosenthal 1979). As a result, it is possible that the results of meta-analyses are skewed towards a more favorable direction. To decrease the possibility of bias, we included both published and unpublished studies. To estimate publication bias in this meta-analysis, we used Egger's linear regression test (Egger et al. 1997) and the Begg and Mazumdar's rank correlation test (Begg and Mazumdar 1994) for each outcome measure. Begg and Mazumdar's rank correlation test was nonsignificant ($p > .05$) for all outcomes except for transfer ($p = .046$). Egger's regression test indicated that all variables were nonsignificant except for retention ($p = .03$) and positive affect ($p = .02$). Further examination of publication bias for retention using Orwin's fail-safe N (Orwin 1983) indicated that an additional 166 studies were needed to nullify the findings, suggesting that publication bias was negligent for retention. In general, the results indicated that publication bias was unlikely for retention, transfer, comprehension, intrinsic motivation, mental effort, perceived difficulty, and

Table 5 Weighted mean effect sizes within change in positive affect

Moderator	N	k	$g+$	SE	95% CI	$Q_{between}$	df	$p_{adjusted}$
Presentation features								
Type of emotional design								
Anthropomorphic shapes	67	1	− 0.02	0.24	[− 0.50, 0.45]			
Anthropomorphic shapes w/ colors	1339	17	0.19**	0.05	[0.08, 0.30]			
Colors	544	5	− 0.21*	0.10	[− 0.42, − 0.01]			
Mixed	107	2	0.08	0.15	[− 0.22, 0.37]			
Between levels (Q_B)						12.03	3	.02
Prior emotion induction								
Yes	528	10	0.24*	0.09	[0.07, 0.41]			
No	1629	15	0.04	0.05	[− 0.06, 0.14]			
Between levels (Q_B)						3.92	1	.11
Nature of presentation								
Animated	1465	20	0.15*	0.06	[0.04, 0.26]			
Static	692	5	− 0.01	0.08	[− 0.16, 0.14]			
Between levels (Q_B)						2.98	1	.12
Pacing of presentation								
Learner-paced	1444	13	0.08	0.09	[− 0.10, 0.26]			
System-paced	636	10	0.24*	0.11	[0.02, 0.46]			
Not reported	77	2	0.10	0.28	[− 0.45, 0.66]			
Between levels (Q_B)						5.42	2	.11
Participant characteristics								
Grade								
Middle school	347	5	0.33*	0.11	[0.12, 0.54]			
High school	67	1	− 0.02	0.24	[− 0.50, 0.45]			
Postsecondary	1743	19	0.05	0.05	[− 0.05, 0.15]			
Between levels (Q_B)						5.77	2	.11
Language/culture								
American	195	4	0.16	0.14	[− 0.12, 0.43]			
Chinese	625	9	0.35**	0.08	[0.19, 0.51]			
Czech	248	3	− 0.09	0.13	[− 0.34, 0.16]			
German	812	7	− 0.05	0.08	[− 0.21, 0.10]			
Other	277	2	0.01	0.12	[− 0.23, 0.24]			
Between levels (Q_B)						16.01	4	.01
Contextual features								
Study settings								
Classroom	277	2	0.01	0.12	[− 0.23, 0.24]			
Laboratory	1880	23	0.11*	0.49	[0.01, 0.21]			
Between levels (Q_B)						0.64	1	.45
Instructional time								
Short	585	9	0.31*	0.08	[0.15, 0.47]			
Moderate	495	6	0.22*	0.09	[0.04, 0.39]			
Long	356	2	− 0.49**	0.15	[− 0.79, − 0.20]			
Not reported	721	8	− 0.02	0.07	[− 0.17, 0.13]			
Between levels (Q_B)						26.52	3	< .001
Learning domain								
Computer science	255	1	− 0.05	0.13	[− 0.30, 0.19]			
Life science	904	12	0.09	0.7	[− 0.04, 0.22]			
Physical science	998	12	0.14*	0.07	[0.004, 0.28]			
Between levels (Q_B)						1.86	2	.45
Publication								
Conference	67	1	− 0.02	0.24	[− 0.50, 0.45]			
Dissertation	332	3	− 0.02	0.11	[− 0.23, 0.20]			
Journal	1758	21	0.12*	0.05	[0.02, 0.22]			
Between levels (Q_B)						1.59	2	.45

Note: *$p < .05$; **$p < .001$

liking/enjoyment. However, the results also indicate the possibility of publication bias for transfer and positive affect, suggesting careful interpretation of the results.

Discussion

This meta-analysis examined the effect of using emotional designs in multimedia learning compared with neutral designs. Findings obtained from this study were similar to those of Brom et al. (2018), suggesting a small to moderate effect in favor of multimedia learning environments that utilize emotional designs. More importantly, the present meta-analysis helps to validate findings by Brom et al. (2018).

Similar to the findings by Brom et al., the distributions for perceived difficulty, comprehension, and liking/enjoyment were not heterogenous; therefore, no further examination of potential moderators for these outcomes was conducted. Subsequent moderator analyses were conducted for the remaining four outcomes: learning outcomes, mental effort, intrinsic motivation, and change in positive affect, since findings showed that the distribution within each outcome was heterogeneous. It is important to note that perhaps due to the difference in our calculation of effect size for change in positive affect compared with Brom et al., our results indicated that change in positive affect was associated with moderate heterogeneity, and therefore, it was included for subsequent moderator analyses. An examination of effect sizes across the categories of moderators suggest that the effectiveness of emotional designs is dependent on outcome measure, presentation features, participant characteristics, and contextual features.

In general, results from our moderator analysis differed markedly from those reported by Brom et al. (2018). For example, aside from age which was a significant predictor for intrinsic motivation, no other moderators were significant in Brom et al.'s meta-analysis. However, in the present meta-analysis, several moderators were found to be significant. There are several explanations for why this may have occurred. Unlike Brom et al. who used a random-effects model in their meta-analysis, we used a fixed-effects model in this present meta-analysis. As mentioned previously, this decision was made as the studies included in the present meta-analysis were conducted by few researchers who utilized similar research designs to examine the effects of emotional designs in multimedia learning. Additionally, several studies served as replication studies with different populations to validate previous findings (see Plass et al. 2014 and Um et al. 2012). Based on this understanding of the studies included in our meta-analysis, we decided that variation in effect sizes may be due to sampling variance (Borenstein et al. 2009). Furthermore, because we updated the pool of studies included in the meta-analysis, we had included more independent experiments in the present meta-analysis (k = 45) than Brom et al. (k = 33). While most of our outcome variables and moderators were coded in a similar fashion as Brom et al., the effect size for change in positive affect was calculated differently in the present meta-analysis. As more studies provided pre- (time 1) and post- (time 2) scores of their positive affective measures instead of a change in affect score, we manually calculated the change in positive affect in this study. This is notably different from how Brom et al. obtained their effect size for positive affect.

Taking these into account, the following paragraphs seek to answer our research questions and delineate how specific variables may have moderated the effectiveness of learning with emotional designs in a multimedia learning environment.

What Are the Effects of Emotional Designs in Multimedia Compared to Neutral Designs?

Findings from this meta-analysis revealed that emotional designs, in comparison to neutral designs in multimedia learning, significantly enhanced learning outcomes (e.g., retention, transfer, and comprehension), intrinsic motivation, liking/enjoyment, mental effort, positive affect outcomes in learners, and reduced perceived difficulty of learning materials. This is not surprising as previous research show that emotional design features such as warm and highly saturated colors (Palmer and Schloss 2010) and anthropomorphic graphics resembling human-like features (Berry and McArthur 1985; Plass et al. 2014) are likely to evoke positive emotions in learners. In addition, positive emotions have been established to enhance affect, particularly learners' intrinsic motivation and their desire to learn (Isen and Reeve 2005; Pekrun 2006). Overall, results from the meta-analysis indicate that emotional designs do enhance various outcomes when they are integrated into multimedia learning materials. However, we suggest that one should exercise caution to ensure that emotional designs are implemented in a coherent manner, with the objective of the learning task at hand (Mayer and Fiorella 2014).

Further support for the relationship between the use of emotional designs is provided by the CATLM, which theorizes that motivation, affect, and metacognitive strategies support the regulation of cognitive processes, effectively influencing learning (Moreno and Mayer 2007). In the following sections, the discussion highlights the potential relationship between motivation, affect, cognitive, and learning outcomes.

How are Emotional Design Effects Moderated by Presentation Features?

The four presentation features coded in this review (type of emotional design, prior emotion induction, nature of presentation, and pacing of presentations) were statistically significant moderators of the effects of using emotional designs. These moderators have been widely researched in emotional design and multimedia learning research and relate to relevant theories in this field.

Type of Emotional Design The type of emotional design is an important moderator since research suggests that the benefits of including such design features on learning outcomes appear to occur indirectly via learners' affect (Dehn & Ven Mulken 2000; Pett and Wilson 1996) and mental effort. The emotional designs used in studies in this meta-analysis were limited to anthropomorphic graphics, colors that were warm, bright and highly saturated, anthropomorphic graphics with colors, or mixed both anthropomorphic graphics and colors. Studies show that using anthropomorphic graphics and warm, bright, highly saturated colors as design features can elicit positive emotions in learners (Berry and McArthur 1985; Palmer and Schloss 2010). However, results from Plass et al.'s (2014) study revealed that the effectiveness of specific design features is more complex. They found that anthropomorphic graphics alone and those with warm colors induced positive emotions in learners, yet warm colors alone did not influence learners' affective states. Findings from previous studies also vary in terms of the effect of the type of emotional designs on mental effort.

Unlike Brom et al. (2018) who did not find type of emotional design to be a significant moderator, results from our meta-analysis provide varying levels of support for type of emotional design for each outcome. Within retention, designs that used either anthropomorphic graphics or colors were associated with higher effect sizes than studies that used a combination of anthropomorphic graphics with colors. Within mental effort and intrinsic motivation, anthropomorphic graphics were significantly more effective than anthropomorphic graphics and color, and colors alone. However, within satisfaction, colors were most effective, while anthropomorphic graphics and colors were significantly less effective.

Although the majority of the studies included anthropomorphic graphics with colors, the mean effect was statistically significant only for retention and change in positive affect outcomes. Since affective outcomes regulate cognitive processes which influence learning performance (CATLM; Moreno and Mayer 2007), it follows that our results show that the same emotional designs that enhanced motivation also enhanced learners' mental effort, and retention. This is an important contribution because it delineates the relationship between the different factors that affect learning and highlights the need to carefully design learning materials that promote better learning. Future research should examine how anthropomorphic graphics with colors may enhance learners' affect and increase their mental effort.

Prior Emotion Induction Since learners' emotional states have a considerable impact on cognitive processes (Moreno and Mayer 2007; Tyng et al. 2017), prior emotion induction is a moderator of interest as well. However, it is still unclear whether the emotions induced prior to a task are sustainable throughout the lesson. Specifically, Um et al. (2012) found that learners' prior positive emotional state facilitated learning through motivation, satisfaction, and mental effort. However, findings from Park et al. (2015) and Plass et al. (2014) did not support this finding. Contrary to Um et al., results from the mental effort data in this meta-analysis do not support the effect of prior emotion induction on overall cognitive and affective outcomes. Instead, the lack of prior emotion induction was associated with a higher mean effect size for mental effort. This finding suggests that emotions that are induced externally (prior to the learning task) may not facilitate learning. In addition, similar to Brom et al. (2018), prior emotion induction was not a significant moderator for any of the other outcomes examined. Overall, results from the meta-analysis show that prior emotion induction has little impact on learning performance via learners' intrinsic motivation, change in positive affect, and mental effort. Future research is needed to elucidate the impact of learners' emotions on their learning experiences. Specifically, it is important to examine how emotions experienced prior to a learning task can facilitate learning outcomes.

Nature of Presentation The nature of presentation, either animated or static, is a potential facilitator of learning. A previous meta-analysis by Berney and Bétrancourt (2016) found that animated graphics were more beneficial than static graphics ($g = 0.23$). Additional research suggests that animated instructional materials may facilitate learning since they aid in the mental visualization of a process, unlike static images that require the learner to piece individual images together (Höffler and Leutner 2007). Thus, learners receiving animated materials may experience less cognitive load. Results from the retention data indicated that animated materials were associated with a larger positive effect; however, nature of presentation was not a significant moderator. On the other hand, results from the intrinsic motivation data revealed that studies with static materials had a greater effect on learners' intrinsic motivation than animated materials. Results from the mental effort data also revealed that studies with static materials had a greater effect on learners' reports of mental effort than animated materials.

With regard to learning outcomes, findings from the present meta-analysis aligned with previous findings and showed that animated materials alone are more beneficial for enhancing learning (Berney and Bétrancourt 2016), and learners were likely to invest more mental effort for static materials than animated materials (Höffler and Leutner 2007). The results also indicated that static materials had a great effect on intrinsic motivation, suggesting that as materials become more cognitively challenging to process, learners must possess greater intrinsic motivation to complete the task. These findings shed light on the importance of the relationship between cognitive load, mental effort, intrinsic motivation, and learning. To some extent, it also provides theoretical support for the CATML assumption that learning is mediated by metacognitive strategies regulating cognitive and affective processes (Moreno and Mayer 2007). To better understand this finding, further research on the relationship between the nature of presentation with different cognitive and affective outcomes is necessary.

Pacing of Presentation In their meta-analysis, Brom et al. did not include pacing as a moderator analysis because the majority of their studies were coded as learner-paced studies (approximately 12%). However, in the present meta-analysis, the addition of new studies increased the number of system-paced studies to approximately 29% of the all included studies. Thus, pacing was examined as a potential moderator. Apart from Mayer and Estrella (2014) and Shangguan et al. (2020) who both examined whether giving participants control over the pace of their learning with emotional designs influenced learning and affective outcomes, few other studies have explicitly examined the impact of emotional designs on learning and affective outcomes when the lesson is either learner-paced or system-paced. Results from the aforementioned studies present conflicting evidence on the effects of learner pacing on learning, cognitive, and affective outcomes. While Mayer and Estrella (2014) found learner-paced materials to positively enhance learning outcomes, this finding was not replicated by Shangguan et al. (2020). However, both studies reported that self-paced learning did not enhance effort, and Shangguan et al. (2020) also found that participants' motivation was not positively influenced.

Results from our moderator analyses indicate that learning with emotional designs did not have an impact on retention, mental effort, and change in positive affect, regardless of pacing of the learning material. However, learning with emotional designs had a higher impact on intrinsic motivation when lessons were learner-paced as compared to system-paced materials. This finding for motivation provides empirical support for an area of research on pacing that previously lacked support, which is both exciting and promising. While more research is needed to examine how this can potentially aid learning performance, the present findings provide preliminary evidence that integrating emotional designs into lessons may be beneficial for enhancing learning when the materials are learner-paced.

How Do Emotional Design Effects Vary with Participant Characteristics?

Grade Level Results from the moderator analysis within retention, mental effort, and intrinsic motivation showed that grade level was a significant moderator. Within the three sets of data, studies conducted with k-12 participants were associated with a larger effect sizes than those conducted with postsecondary participants. This finding aligns with that by Brom et al. (2018) who found specifically that the effects of emotional designs were stronger for younger participants than college-aged participants. However, it is important to note that since the majority of studies included in this meta-analysis were conducted with postsecondary participants, we approach the interpretation of the results with caution. Furthermore, given our understanding that metacognitive strategies and self-regulation of learning and attention develop with age (Napolitano and Sloutsky 2004; Raffaelli, Crockett and Shen, 2005), additional research examining the effects of grade level on the effects of emotional designs on affective and cognitive outcomes is warranted.

Language/Culture In this meta-analysis, language/culture was also included as a potential moderator to examine how the effectiveness of emotional designs differed across different samples. Language/culture was a statistically significant moderator within intrinsic motivation and change in positive affect outcomes. Specifically, for intrinsic motivation, studies conducted with American and German participants were associated with larger effect sizes than the other languages/cultures for intrinsic motivation. However, for change in positive affect, studies conducted with Chinese participants were associated with a larger effect size. Given the lack of cross-cultural research on the use of emotional designs in multimedia learning, we hesitate to make broad generalizations on the advantages of emotional designs on affective and cognitive outcomes for specific languages/cultures. Yet, the differences across cultures are not entirely surprising, as research indicates cross-cultural differences in motivation and emotion perception between Western and Asian cultures. For example, Grant and Dweck (2001) found that Asian students tend to emphasize effort and performance goals, while Western students tend to focus on mastery goals. Additionally, research on emotion judgment found differences between American and Japanese participants' ratings on emotion intensity (Matsumoto et al. 2002). Since individual differences influence the efficiency of learning with media (Park et al. 2014), we recommend further research examining the impact of language/culture on the effectiveness of emotional designs. Specifically, it would be interesting to consider how language and cultural perception of emotional designs contribute to learners' motivation and investment of effort in learning.

How Are Emotional Design Effects Moderated by Contextual Features of Studies?

Instructional Time The amount of instructional time provided to learners is important for them to comprehend the learning material and meet the learning objective (Mayer 2014). A reasonable assumption is that extended exposure to learning materials may increase learning performance and foster positive attitudes towards learning. However, results from the mental effort analysis indicated that moderate instruction time (between 15 and 30 min) had a significantly higher effect size than studies conducted with a moderate or long instruction time. Furthermore, results from the change in positive affect data showed that studies of short and moderate lengths of instruction time (less than 15 min and between 15 and 30 min) were associated with a higher mean effect than studies with long instruction times. It may be possible that studies with moderate instructional time had slightly more detailed or complex content than studies with shorter instructional time. For this reason, it is plausible that emotional designs can have a higher impact on participants' mental effort when the lesson is moderately long because students have sufficient time to focus on the task at hand to make sense of things. With regard to change in positive affect, short to moderate exposure to emotional designs might have a higher impact on participants' positive affect than long instructional time. Perhaps, the novelty of learning with emotional designs was most beneficial between short and moderate lengths of instructional time. Yet, with these plausible explanations for findings for mental effort and change in positive affect, it is still unclear why studies longer than 30 min had a smaller effect size than studies between 15 and 30 min. Ideally, the longer exposure should increase familiarity with the lesson and foster more positive emotions and attitudes towards learning. Future research should certainly consider how and why the effectiveness of emotional designs wanes over time.

Learning Domain Results from the moderator analysis revealed that learning domain was a significant moderator for retention and intrinsic motivation. Within retention, studies coded as others were associated with a large effect size; however, studies in life science and physical science were also associated with small positive effect sizes. Within intrinsic motivation, both studies in life science and studies coded as others were associated with small positive effect sizes. While this suggests that emotional designs may be effective for enhancing retention and intrinsic motivation in specific domains over others, we also recognize that majority of the studies were conducted in the sciences and excluded the arts and social sciences domains. Therefore, we refrain from making any generalizations about this finding. Nevertheless, we encourage future research to broaden the scope of this research to other learning domains.

Limitations and Future Research Directions

This meta-analysis revealed evidence for the moderating factors of retention, mental effort, intrinsic motivation, and change in positive affect with emotional designs (specifically the use of colors and anthropomorphic graphics) in multimedia learning. However, this meta-analysis is not without any limitations. This study sought to replicate and extend the previous meta-analysis by Brom et al. (2018). However, the overall number of studies included in the meta-analysis is still relatively small. We were also limited to the information provided in research studies for change in positive affect. Whereas Brom et al. utilized the change measure for positive affect, in this meta-analysis, we calculated change in positive affect from the reported pre and post positive affect scores as studies more frequently reported pre and post affect scores than change scores. For this reason, our findings for change in positive affect differed from that of Brom et al.

While this meta-analysis is not the first review on the effects of emotional designs in multimedia learning (see Brom et al. 2018), it supports the previous study by highlighting the need for more research to better understand the factors that influence learning with emotional designs. Specifically, more studies are needed to elucidate the effects of participant characteristics and contextual features on learning with emotional designs. For example, a common criticism of multimedia learning research is that treatment interventions are too short to represent ecologically valid educational contexts (Mayer and Estrella 2014; Mayer 2017). Hence, future studies should examine the effectiveness of using emotional designs in other instructional domains, in ecologically valid environments and with longer instructional time. Finally, future research on emotional designs in multimedia should also be conducted with different populations of students. Due to convenience samples available on college campuses, majority of the studies included in the present meta-analysis were conducted with undergraduate students. To increase our understanding of how emotional designs may affect learners of all grades and cultures, future studies should explore the effects of emotional designs on k-12 students and a more diverse population of participants.

Conclusion

The purpose of the present meta-analysis was to serve as a replication and extension of Brom et al.'s (2018) meta-analysis. The findings from this study appear to validate the findings obtained in the previous meta-analysis. Like Brom et al.'s meta-analysis, the findings from this present meta-analysis have several theoretical and practical implications for research on multimedia learning and the implementation of multimedia learning materials. The results provide empirical support for Moreno's Cognitive Affective Theory of Learning with Media (Moreno 2006) and the emotions-as-facilitator-of-learning hypothesis (Um et al. 2012). Although affective factors have been overlooked in multimedia learning research, the introduction of the CATML has inspired researchers to consider how both cognitive and affective factors influence learning with multimedia. Findings from this meta-analysis demonstrate that emotional designs increase the motivational and affective outcomes that help mediate cognitive engagement for learning (Plass & Kalyuga 2019). Specifically, learning performance may be mediated by motivational and affective outcomes. This also suggests that appropriate use of emotional designs can enhance the affective outcomes that regulate cognitive outcomes and the learning experience.

While results show preliminary support for the continued use of emotional designs in learning, further research is needed in non-STEM domains, as well as with a greater diversity of students in terms of age and cultural background. This will deepen researchers' understanding of the moderators that impact the effects of emotional designs. The use of emotional designs appears to be particularly beneficial for students in grades k-12. Instructional designers and educators are encouraged to consider appropriate ways to implement emotional designs within multimedia learning environments to enhance students' learning and affective outcomes. However, we caution against providing designs that are not aligned with the learning objectives, as those could include seductive details that distract students and have a deleterious effect on learning (Wang and Adesope 2017; Harp and Mayer 1998; Rey 2012). Overall, findings from this meta-analysis can guide educators and instructional designers in implementing emotional designs within multimedia learning environments to enhance students' affect, which could enhance learning outcomes.

Compliance with Ethical Standards

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

References

The asterisk indicates studies included in this meta-analysis.

Baddeley, A. (1992). Working memory. *Science, 255*(5044), 556–559. https://doi.org/10.1126/science.1736359.

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics, 50*(4), 1088–1101. https://doi.org/10.2307/2533446.

Benjaminni, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B: Methodological, 57*(1), 289–300.
Berney, S., & Bétrancourt, M. (2016). Does animation enhance learning? A meta-analysis. *Computers & Education, 101*, 150–167. https://doi.org/10.1016/j.compedu.2016.06.005.
Berry, D. S., & McArthur, L. Z. (1985). Some components and consequences of a babyface. *Journal of Personality and Social Psychology, 48*(2), 312–323. https://doi.org/10.1037/0022-3514.48.2.312.
Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, England: Wiley.
Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. R. (2013). *Comprehensive meta-analysis* (Version 3.0) [Computer software]. Englewood, NJ.
Brand, S., Reimer, T., & Opwis, K. (2007). How do we learn in a negative mood? Effects of negative mood on transfer and learning. *Learning and Instruction, 17*(1), 1–16. https://doi.org/10.1016/j.learninstruc.2006.11.002.
*Brom, C., Hannemann, T., Stárková, T., Bromová, E., & Děchtěrenko, F. (2016). Anthropomorphic faces and funny graphics in an instructional animation may improve superficial rather than deep learning: a quasi-experimental study. In J. Novotna & A. Jancarik (Eds.), *Proceedings of the 15th European Conference on e-Learning,* Prague, Czech Republic (pp. 89-97). Academic Conference and Publishing International Limited: Reading, UK.
Brom, C., Stárková, T., & D'Mello, S. K. (2018). How effective is emotional design? A meta-analysis on facial anthropomorphisms and pleasant colors during multimedia learning. *Educational Research Review, 25*, 100–119. https://doi.org/10.1016/j.edurev.2018.09.004.
*Chui, T. K., Jong, M. S., & Mok, I. A. C. (2020). Does learner expertise matter when designing emotional multimedia for learners of primary school mathematics? *Educational Technology Research and Development.* Advanced online publication. https://doi.org/10.1007/s11423-020-09775-4
Dehn, D. M., & van Mulken, S. (2000). The impact of animated interface agents: a review of empirical research. *International Journal of Human-Computer Studies, 52*(1), 1–22. https://doi.org/10.1006/ijhc.1999.0325.
*Dong, C. (2007). *Positive emotions and learning: what makes a difference in multimedia design?* (Unpublished doctoral dissertation). New York University: New York.
Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple graphical test. *British Medical Journal, 315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629.
Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: unlocking the file drawer. *Science, 345*(6203), 1502–1505. https://doi.org/10.1126/science.1255484.
*Gong, S., Shangguan, C., Zhai, K., & Guo, Y. (2017). The effects of emotional design on multimedia learning. *Acta Psychologica Sinica*, *49*(6), 771-782. https://doi.org/10.3724/SP.J.1041.2017.00771
Gorn, G. J., Chattopadhyay, A., Yi, T., & Dahl, D. W. (1997). Effects of color as an executional cue in advertising: they're in the shade. *Management Science, 43*(10), 1387–1400. https://doi.org/10.1287/mnsc.43.10.1387.
Grant, H., & Dweck, C. S. (2001). Cross-cultural response to failure: considering outcome attributions with different goals. In F. Salili, G. Ghiu, & Y. Hong (Eds.), *Student motivation: the culture and context of learning* (pp. 203–219). New York: Kluwer Academic/Plenumm Publishers.
*Haaranen, L., Ihantola, P., Sorva, J., & Vihavainen, A. (2015). In search of the emotional design effect in programming. In *Proceedings of the 37th International Conference on Software Engineering,* ICSE 2015, *2*, 428-434.
Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: a theory of cognitive interest in science learning. *Journal of Educational Psychology, 50*, 414–434. https://doi.org/10.1037/0022-0663.90.3.414.
Hasler, B. S., Kersten, B., & Sweller, J. (2007). Learner control, cognitive load and instructional animation. *Applied Cognitive Psychology, 21*(6), 713–729. https://doi.org/10.1002/acp.1345.
Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6*(2), 107–128.
*Heidig, S., Müller, J., & Reichelt, M. (2015). Emotional design in multimedia learning: differentiation on relevant design features and their effects on emotions and learning. *Computers in Human Behavior*, *44*, 81–95. https://doi.org/10.1016/j.chb.2014.11.009
Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*(11), 1539–1558.
Höffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: a meta-analysis. *Learning and Instruction, 17*(6), 722–738. https://doi.org/10.1016/j.learninstruc.2007.09.013.
Höffler, T. N., & Schwartz, R. N. (2011). Effects of pacing and cognitive style across dynamic and non-dynamic representations. *Computers & Education, 57*(2), 1716–1726. https://doi.org/10.1016/j.compedu.2011.03.012.

Isen, A., & Reeve, M. (2005). The influence of positive affect on intrinsic and extrinsic motivation: facilitating enjoyment of play, responsible work behavior, and self-control. *Motivation and Emotion, 29*(4), 295–323. https://doi.org/10.1007/s11031-006-9019-8.

Isen, A. M., Daubman, K. A., & Nowicki, G. P. (1987). Positive affect facilitates creative problem solving. *Journal of Personality and Social Psychology, 52*(6), 1122–1131. https://doi.org/10.1037/0022-3514.52.6.1122.

Jamet, E., Gavota, M., & Quaireau, C. (2008). Attention guiding in multimedia learning. *Learning and Instruction, 18*(2), 135–145. https://doi.org/10.1016/j.learninstruc.2007.01.011.

Kumar, J. A. (2016). *Effects of multimedia induced emotions on achievement, intrinsic motivation and satisfaction among polytechnic students with different levels of emotional intelligence.* (Unpublished doctoral dissertation). Malaysia: Universiti Sains Malaysia.

*Kumar, J. A., Balakrishnan, M., & Yahaya, W. A. J. W. (2016). Emotional design in multimedia learning: how emotional intelligence moderates learning outcomes. *International Journal of Modern Education and Computer Science*, *8*(5), 54–63. https://doi.org/10.5815/ijmecs.2016.05.07

*Kumar, J. A., Balakrishnan, M., & Yahaya, W. A. J. W. (2019). Exploring the effects of emotional design and emotional intelligence in multimedia-based learning: an engineering educational perspective. *New Review of Hypermedia and Multimedia*, *25*(1-2), 57-86. https://doi.org/10.1080/13614568.2019.1596169

Lakens, D., Hilgard, J., & Staaks, J. (2016). On the reproducibility of meta-analyses: six practical recommendations. *BMC psychology, 4*(1), 1–10. https://doi.org/10.1186/s40359-016-0126-3.

*Le, Y., Liu, J., Deng, C., & Dai., D. Y. (2018) Heart rate variability reflects the effects of emotional design principle on mental effort in multimedia learning. *Computers in Human Behavior*, *89*, 40-47. https://doi.org/10.1016/j.chb.2018.07.037

Linnenbrink, E. A. (2006). Emotion research in education: Theoretical and methodological perspectives on the integration of affect, motivation, and cognition. *Educational Psychology Review, 18*(4), 307–314. https://doi.org/10.1007/s10648-006-9028-x .

Lowe, R. K., & Schnotz, W. (2014). Animation principles in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., p. 513–546). New York, NY: Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.026

Matsumoto, D., Consolacion, T., Yamada, H., Suzuki, R., Franklin, B., Paul, S., et al. (2002). *Cognition and Emotion, 16*(6), 721–747. https://doi.org/10.1080/02699930143000608.

Mayer, R. E. (2014). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 43–71). New York, NY: Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.005.

Mayer, R. E. (2017). Using multimedia for e-learning: multimedia for e-learning. *Journal of Computer Assisted Learning, 33*(5), 403–423. https://doi.org/10.1111/jcal.12197.

*Mayer, R. E., & Estrella, G. (2014). Benefits of emotional design in multimedia instruction. *Learning and Instruction*, *33*, 12–18. https://doi.org/10.1016/j.learninstruc.2014.02.004

Mayer, R. E., & Fiorella, L. (2014). Principles for reducing extraneous processing in multimedia learning: coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 279–315). New York, NY: Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.015.

Mega, C., Ronconi, L., & De Beni, R. (2014). What makes a good student? How emotions, self-regulated learning, and motivation contribute to academic achievement. *Journal of Educational Psychology, 106*(1), 121–131. https://doi.org/10.1037/a0033546.

*Miller, C. (2011). Aesthetics and e-assessment: the interplay of emotional design and learner performance. *Distance Education*, *32*(3), 307–337. https://doi.org/10.1080/01587919.2011.610291

*Moremoholo, T. P., & de Lange, R. W. (2018). Anthropomorphic graphics: how useful are they as an instructional aid to facilitate learning? *The Independent Journal of Teaching and Learning*, *13*(2), 67–81.

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*(3), 149–158.

Moreno, R., & Mayer, R. E. (2007). Interactive multimodal learning environments. *Educational Psychology Review, 19*(3), 309–326.

Münchow, H. (2017). *I feel therefore I learn – effectiveness of affect induction interventions and possible covariates on learning outcomes.* (Unpublished doctoral dissertation). Würzburg: Julius-Maximillians-Universität: Würzburg, Germany.

*Münchow, H., & Bannert, M. (2019). Feeling good, learning better? Effectivity of an emotional design procedure in multimedia learning. *An International Journal of Experimental Educational Psychology*, *39*(4), 530-549. https://doi.org/10.1080/01443410.2018.1524852

*Münchow, H., Mengelkamp, C., & Bannert, M. (2017). The better you feel the better you learn: do warm colors and rounded shapes enhance learning outcome in multimedia learning? *Education Research International*, *2017*, 1-15. doi:https://doi.org/10.1155/2017/2148139
Napolitano, A. C., & Sloutsky, V. M. (2004). Is a picture worth a thousand words? The flexible nature of modality dominance in young children. *Child Development, 75*(6), 1850–1870. https://doi.org/10.1111/j.1467-8624.2004.00821.x.
*Navratil, S. D., Kühl, T., & Heidig, S. (2018). Why the cells look like that – the influence of learning with emotional design and elaborative interrogations. *Frontiers in Psychology*, *9*, 1–15. https://doi.org/10.3389/fpsyg.2018.01653.
*Ng, K. H., & Chiu, T. K. F. (2017). Emotional multimedia design for developing mathematical problem-solving skills. In W. W. K. Ma, C.-K. Chan, K. Tong, H. Fung, & C. W. R. Fong (Eds.), *New ecology for education — communication x learning* (pp. 131–141). Singapore: Springer Singapore. https://doi.org/10.1007/978-981-10-4346-8_11
*Nurminen, M. (2016). Effects of emotional design and goal orientation on students' emotions and learning outcomes in university programming education (Unpublished thesis). Tampere University of Technology: Tampere, Finland.
*Nurminen, M. I., Leppanen, L., Vaataja, H., & Ihantola, P. (2017). Search of the emotional design effect in programming revised. In É. Lavoué, H. Drachsler, K. Verbert, J. Broisin, & M. Pérez-Sanagustín (Eds.), *Data driven approaches in digital education* (Vol. 10474). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-66610-5
Orwin, R. G. (1983). A Fail-safe N for effect size in meta-analysis. *Journal of Educational Statistics, 8*(2), 157–159. https://doi.org/10.2307/1164923.
Paivio, A. (1986). *Mental representations: a dual-coding approach*. Oxford, England: Oxford University Press.
Palmer, S. E., & Schloss, K. B. (2010). An ecological valence theory of human color preference. *Proceedings of the National Academy of Sciences, 107*(19), 8877–8882. https://doi.org/10.1073/pnas.0906172107.
Park, B., Plass, J. L., & Brünken, R. (2014). Cognitive and affective processes in multimedia learning. *Learning and Instruction, 29*, 125–127. https://doi.org/10.1016/j.learninstruc.2013.05.005.
*Park, B., Knörzer, L., Plass, J. L., & Brünken, R. (2015). Emotional design and positive emotions in multimedia learning: an eyetracking study on the use of anthropomorphisms. *Computers & Education*, *86*, 30–42. https://doi.org/10.1016/j.compedu.2015.02.016
Pekrun, R. (2006). The control-value theory of achievement emotions: assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), 315–341. https://doi.org/10.1007/s10648-006-9029-9.
Pett, D., & Wilson, T. (1996). Color research and its application to the design of instructional materials. *Educational Technology Research and Development, 44*(3), 19–35. https://doi.org/10.1007/BF02300423.
Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology, 95*(4), 667–686. https://doi.org/10.1037/0022-0663.95.4.667.
Plass, J. L., & Kalyuga, S. (2019). Four ways of considering emotion in cognitive load theory. *Educational Psychology Review, 31*(2), 339–359. https://doi.org/10.1007/s10648-019-09473-5.
Plass, J. L., & Kaplan, U. (2016). Emotional design in digital media for learning. In S. Y. Tettegah & M. Gartmeier (Eds.), *Emotions, technology, design, and learning* (pp. 131–161). New York: Elsevier.
*Plass, J. L., Heidig, S., Hayward, E. O., Homer, B. D., & Um, E. (2014). Emotional design in multimedia learning: effects of shape and colour on affect and learning. *Learning and Instruction*, *29*, 128–140. https://doi.org/10.1016/j.learninstruc.2013.02.006
Raffaelli, M., Crockett, L. J., & Shen, Y.-L. (2005). Developmental stability and change in self-regulation from childhood to adolescence. *Journal of Genetic Psychology, 166*(1), 54–76.
Rey, G. D. (2012). A review of research and a meta-analysis of the seductive detail effect. *Educational Research Review, 7*(3), 216–237. https://doi.org/10.1016/j.edurev.2012.05.003.
Rop, G., Schüler, A., Verkoeijen, P. P. J. L., Scheiter, K., & Van Gog, T. (2018). Effects of task experience and layout on learning from text and pictures with or without unnecessary picture descriptions. *Journal of Computer Assisted Learning, 34*(4), 458–470. https://doi.org/10.1111/jcal.12287.
Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin, 86*(3), 638–641. https://doi.org/10.1037/0033-2909.86.3.638.
*Schneider, S., Häßler, A., Habermeyer, T., Beege, M., & Rey, G. D. (2018a). The more human, the higher the performance? Examining the effects of anthropomorphism on learning with media. *Journal of Educational Psychology, 111*(1), 57–72. https://doi.org/10.1037/edu0000273.
*Schneider, S., Nebel, S., Beege, M., & Rey, G. D., (2018b). Anthropomorphism in decorative pictures: Benefit or harm for learning? *Journal of Educational Psychology*, *110*(2), 218–232. https://doi.org/10.1037/edu0000207 .

*Shangguan, C., Wang, Z., Gong, S., Guo, Y., & Xu, S. (2019). More attractive or more interactive? The effects of multi-leveled emotional design on middle school students' multimedia learning. *Frontiers in Psychology*, *10*, 1-12. https://doi.org/10.3389/fpsyg.2019.03065.

*Shangguan, C., Gong, S., Guo, Y., Wang, X., & Lu, J. (2020). The effects of emotional design on middle school students' multimedia learning: the role of learners' prior knowledge. *Educational Psychology*, 1–18. https://doi.org/10.1080/01443410.2020.1714548

*Stárková, T., Lukavský, J., Javora, O., & Brom, C. (2018). Anthropomorphisms in multimedia learning: attract attention but do not enhance learning? *Journal of Computer Assisted Learning*, *35*, 555-568. https://doi.org/10.1111/jcal.12359

Tabbers, H. K., & de Koeijer, B. (2010). Learner control in animated multimedia instructions. *Instructional Science, 38*(5), 441–453. https://doi.org/10.1007/s11251-009-9119-4.

Tabbers, H. K., Martens, R. L., & Merriënboer, J. J. G. (2004). Multimedia instructions and cognitive load theory: effects of modality and cueing. *British Journal of Educational Psychology, 74*(1), 71–81. https://doi.org/10.1348/000709904322848824.

Tyng, C. M., Amin, H. U., Saad, M. N. M., & Malik, A. S. (2017). The influences of emotion on learning and memory. *Frontiers in Psychology, 8*, 1–22. https://doi.org/10.3389/fpsyg.2017.01454.

Um, E. R. (2008). *The effect of positive emotions on cognitive processes in multimedia -based learning*. (Publication No. 3308294), [Doctoral Dissertation, New York University]. ProQuest Dissertations Publishing.

*Um, E. R., Song, H. S., & Plass, J. L. (2007). The effect of positive emotions on multimedia learning. In C. Montgomerie & J. Seale (Eds.), *EdMedia: World Conference on Educational Media and Technology*, Vancouver, Canada (pp. 4176–4185). Association for the Advancement of Computing in Education (AACE): Waynesville, NC.

*Um, E. R., Plass, J. L., Hayward, E. O., & Homer, B. D. (2012). Emotional design in multimedia learning. *Journal of Educational Psychology*, *104*(2), 485–498. https://doi.org/10.1037/a0026609 .

*Uzun, A. M., & Yildirim, Z. (2018). Exploring the effect of using different levels of emotional design features in multimedia science learning. *Computers & Education*, *119*, 112-128. https://doi.org/10.1016/j.compedu.2018.01.002

Wang, Z., & Adesope, O. (2017). Do focused self-explanation prompts overcome seductive details? A multimedia study. *Journal of Educational Technology & Society*, *20*(4), 47–57.

Waytz, A., Cacioppo, J., & Epley, N. (2010). Who sees human? The stability and importance of individual differences in anthropomorphism. *Perspectives on Psychological Science, 5*(3), 219–232. https://doi.org/10.1177/1745691610369336.

Wittrock, M. C. (1989). Generative processes of comprehension. *Educational Psychologist, 24*(4), 345–376. https://doi.org/10.1207/s15326985ep2404_2.

Wolfson, S., & Case, G. (2000). The effects of sound and color on responses to a computer game. *Interacting with Computers*, *13*(2), 183–192. https://doi.org/10.1016/S0953-5438(00)00037-0