# Unintended Emotions in the Laboratory: Emotions Incidentally Induced by a Standard Visual Working Memory Task Relate to Task Performance

Sara Laybourn, Anne C. Frenzel, Martin Constant, and Heinrich R. Liesefeld
Department Psychologie, Ludwig-Maximilians-Universität, München, Germany

## Abstract

The ability to temporarily hold information in visual working memory (VWM) is among the most crucial and most extensively examined human cognitive functions. Here, we empirically confirm previous speculations (1) that a standard VWM task arouses emotions in participants and (2) that these task-induced emotions are related to VWM performance. In a first qualitative study ($N = 19$), by adapting a qualitative method of inquiry, the think-aloud technique, we found that the task induced different positive and negative emotions, such as joy and anger, which varied on the inter- as well as on the intra-individual level. The emotional experiences seemed to be tied to the implicit achievement requirement of the VWM task (getting it right vs. wrong). Encouraged by these findings, two quantitative studies ($N = 45$, and $N = 44$, respectively) revealed that VWM performance was positively linked to joy and pride, and negatively linked to anger, frustration and boredom on the inter- and on the intra-individual level. Notably, these emotions were also affected by an experimental manipulation of task difficulty (set-size 4 vs. 8). Further, the findings from Study 3 were replicated in a fourth high-powered online study ($N = 110$). This research is the first to demonstrate that a task designed to measure VWM in itself triggers emotions, specifically achievement emotions, which, in turn, are linked with VWM performance. Our findings suggest that these task-induced emotions should be considered as potential confounding variables in future research on VWM and in cognitive research in general.

*Keywords:* visual working memory, achievement emotions, task-induced emotions

Visual working memory (VWM), which can be defined as "the active maintenance of visual information to serve the needs of ongoing tasks" (Luck & Vogel, 2013, p. 392), is crucial in everyday life (Alloway & Alloway, 2010; Conway et al., 2003; Fukuda et al., 2010). Researchers have developed computer-based VWM tasks, such as the change detection task (Luck & Vogel, 1997; Pashler, 1988) or the continuous color-report task (Wilken & Ma, 2004; Zhang & Luck, 2008), which have been shown to provide reliable and valid VWM estimates of individuals' VWM capacities (Johnson et al., 2013; Kyllingsbæk & Bundesen, 2009). Studies using these tasks have brought about valuable insights into the basic functioning of human visual memory (Luck & Vogel, 2013). However, this research tradition is dominated by experimental designs where (typically small) sample means are compared across experimental conditions, and any inter- and intra-individual differences are typically considered as "noise" (Kanai & Rees, 2011; Vogel & Awh, 2008). In the

present contribution, we propose that participants and their complexity may have been oversimplified in such existing research paradigms. Although oversimplification is a crucial ingredient of research in general, looking into the "noise" can also be fruitful, and it can even be crucial if it proves to be confounded with experimental manipulations of interest or if it affects the validity of the measures. Specifically, we propose to take into consideration one important human factor: the emotions participants feel while performing experimental laboratory tasks (see also Dukes et al., 2021). We sense that typical prior research seems to view participants as machines who enter the laboratory and perform VWM tasks as successfully as their "hardware" (i.e., their VWM capacity) allows, while neglecting the task-induced emotional experiences. To the authors' knowledge, there is currently no research on how participants feel during VWM tasks in the laboratory, that is the emotions which emerge *as a result of performing the task itself*, and whether these emotions, in turn, are systematically

linked with VWM performance, thus potentially biasing capacity estimates.

There are some hints in the literature that emotions induced by the VWM task may be linked to individual differences in VWM performance (Luck, 2014; Rouder et al., 2008). Those are predominantly informal observations and largely speculative, which, to our knowledge, have not been researched systematically. There is substantial empirical evidence, however, that externally induced emotional states influence participants' VWM performance (Spachtholz et al., 2014; Xie and Zhang, 2016, but see Souza et al., 2021). Therefore, it is likely that, to the degree that emotional experiences occur as a result of performing the VWM task, they may well be linked with VWM performance.

The aim of the present research was twofold: (1) to establish whether participants experience emotions during a VWM task, which are induced by the task itself, and (2) if this was the case, to explore how these incidentally induced emotions are related to VWM performance.

To answer these research questions, we followed a mixed-method approach, by first conducting a qualitative study using an approach based on the think aloud method (van Someren et al., 1995) to investigate whether participants experienced any task-induced emotions during a VWM task in the laboratory, and if so, which emotions those were (some authors also refer to this as "emote-aloud," see e.g., D'Mello and Graesser, 2012). In three further quantitative studies, we explored the link between self-reported discrete task-induced emotions and VWM performance, from an inter-individual and intra-individual perspective. Our reasoning for these studies is outlined in the following.

---

Sara Laybourn

Anne C. Frenzel

Martin Constant

Heinrich R. Liesefeld

Martin Constant and Heinrich R. Liesefeld are now at the Department of Psychology, University of Bremen, Germany

Correspondence concerning this article should be addressed to Sara Laybourn or Anne C. Frenzel, Department Psychologie, Ludwig-Maximilians-Universität, Leopoldstr. 13, 80802 Munich, Germany, Email: sara.laybourn@psy.lmu.de or frenzel@lmu.de

## Emotions and Visual Working Memory

Emotions are generally defined as multi-component processes, including affective, motivational, cognitive, physiological, and expressive components (Mulligan & Scherer, 2012), which can influence our thoughts, behaviors, and performance (Beal et al., 2005). Research has shown that emotions also influence higher cognitive functions, such as memory. In research regarding long-term memory, it is well established that the emotions experienced at the time of memory formation influence how the memory is stored (e.g., LaBar and Cabeza, 2006). In addition, there is also a specific body of literature which explored the effects of emotions on working memory, specifically VWM. However, these findings are somewhat ambiguous, as we briefly outline below.

When studying the effect of emotional states on VWM, negative, positive and/or neutral emotional states are typically induced prior to the VWM task. In doing so, some researchers have found evidence that negative emotional states relate to poorer working memory performance. For instance, Figueira et al. (2017) showed that participants' contralateral delay activity (CDA) amplitudes were significantly lower in a negative emotional condition as compared to a neutral emotional condition. CDA is an electrophysiological index of VWM processing, which has been shown to correlate with the amount of information stored in working memory (Luria et al., 2016; Vogel & Machizawa, 2004). Further, on an intra-individual level, Brose et al. (2014) found that daily variations in participants' mood (measured over a 100 day period) related to working memory performance. Specifically, negative affect was negatively linked with spatial working memory performance. The authors also found that variations in positive affect were positively linked to working memory performance.

Digging deeper into the potential effects of emotional states on working memory performance, some authors proposed to explore potentially differential effects on the quality versus quantity of visual performance. For instance, Spachtholz et al. (2014) randomly assigned participants to a condition inducing either neutral or negative emotions prior to performing the continuous color-report task. Results showed that the number of remembered items was lower in the negative-emotion condition but VWM performance in terms of quality (precision of color memory) increased. The authors concluded that emotional state leads to a tradeoff between quantity and quality in terms of VWM performance, with negative emotional states favoring quality over quantity. However, those effects were not fully confirmed by Xie and Zhang (2016): manipulating emotional state on a trial by trial basis via presentation of IAPS pictures (International Affective Picture System, Lang et al., 1997), they found that self-reported negative emotion yields higher VWM precision, but this time without any concurrent decrease in VWM capacity (see also Long et al., 2020, who found such a tradeoff). How-

ever, a large-scale attempt to replicate those findings failed to find any evidence of an enhancing effect of self-reported negative emotion on either precision or capacity with seven experiments in three different labs and with participants from four different countries (Souza et al., 2021).

Overall, when considering the results discussed above, it seems there is some indication for the relevance of emotional states in VWM performance. However, the results are not straightforward and attempts to replicate the findings yield even more ambiguous results. Importantly, this existing research rests mainly on studies which induced emotional states prior to the VWM task (either block-/session-wise or trial-wise). Not only may such emotion induction procedures lack ecological validity, but more importantly, they are ineffective in exploring another, potentially more relevant factor: the emotions participants experience because of the VWM task itself. These may be systematically linked with performance, and may also differ across experimental conditions, thus creating an important potential confound and posing a threat to the validity of the VWM performance scores. To the authors' knowledge, no study to date has explored these assumptions.

Despite the eminent lack of systematic research into potential links between incidentally induced emotions and VWM performance scores, some researchers from the VWM community hint towards the possibility that tasks designed to measure cognitive functions may induce emotions in participants, and that these emotions in turn may be linked to participants' performance. For instance, in his book on event-related potentials (ERPs), Luck (2014) states the following: "ERP experiments tend to be long and boring, with trial after trial of the same basic task. To ensure that you are collecting the highest quality data possible, it is important to keep your subjects happy and relaxed. If they are unmotivated or become bored, they may not pay close attention to their performance, weakening your effects" (p. 144). Also the classical VWM tasks are typically lengthy, requiring participants to go through multiple-trial-blocks of either change detection or active color recall, which typically take up to 45 minutes or more to complete. To the authors' knowledge, whether or not these tasks are generally not enjoyable or evoke boredom in participants because they are long and repetitive is based purely on introspection or speculation, and has never been explored systematically. Moreover, the presumption that such task-induced emotions may also have an effect on the obtained results has not been tested.

Further, researchers in the field of VWM have conveyed their thoughts on possible task-induced emotions and how these emotions may be linked to participants' performance. For instance, Rouder and colleagues (2008) found that some participants performed worse than expected on difficult as opposed to easy trials of a VWM task, and they concluded that this may have been the case because some participants were "intimidated" by the difficult trials (p. 5978). Similarly, Spachtholz et al. (2014) speculated that differences in VWM performance may "be brought about unintentionally by cues such as affective state that signal requirements of the current situation" (p. 1455).

Overall, it seems highly likely that engaging in a VWM task triggers emotional experiences in participants. It has been speculated that they may be perceived as boring or intimidating, yet we argue they may also be experienced as challenging and engaging. The first goal of the present research was therefore to explore what participants feel when participating in a typical VWM task.

Our second goal was to explore the links between emotions induced by a VWM task and VWM performance. If prior speculations are correct and emotional experiences occur because of certain characteristics of the VWM task (i.e., difficult trials, see Rouder et al., 2008), these emotions may be potentially confounding variables distorting the VWM performance estimates.

Evidence from more applied, educational psychology has demonstrated that engaging in tasks can trigger emotional experiences, which in turn are linked with performance. A brief overview of these results is given in the following.

## Task-induced Emotions and their Link with Performance

Emotions are thought to be activated by individual appraisals of specific objects or events (Mulligan & Scherer, 2012). Tasks and activities can also act as objects, which trigger emotional experiences. For instance, people experience joy when engaging in an activity that they appraise as pleasant, either because they receive an extrinsic reward for their engagement in the task, such as praise, or because the task itself is rewarding to the individual (Csikszentmihalyi, 2000). Specifically, task enjoyment is thought to play a major role in different concepts such as flow experience (Csikszentmihalyi, 2000), intrinsic motivation (Deci & Ryan, 1991), and achievement motivation (Dweck & Elliott, 1983) – concepts which are important for performance.

The role of task-induced emotions in cognitive performance is commonly studied in academic settings. Here, the research focus lies mainly on achievement or epistemic emotions. Achievement emotions can either relate to achievement outcomes, such as exam results, or to achievement-related activities, such as studying or class participation (Pekrun et al., 2011). By definition, achievement emotions emerge as a result of achievement outcomes, that is, success entails positive emotions, and failure entails negative emotions. In addition, achievement emotions can be assumed to influence learners' cognitive resources, motivation, strategy use, and self-regulated learning such that they in turn predict achievement outcomes (Goetz & Hall, 2013). Research on achievement emotions has shown consistently that negative achievement emotions, such as shame and anger,

are linked negatively with performance (e.g., Pekrun and Perry, 2014) whereas positive achievement emotions, such as joy and pride, correlate positively with performance (e.g., Pekrun, Lichtenfeld, et al., 2017). In longitudinal research designs, it has been shown that those achievement emotion—performance links are driven by reciprocal causation, with emotions and performance predicting each other over time (Pekrun, Lichtenfeld, et al., 2017; Putwain et al., 2018).

The cognitive characteristics of a task can also induce emotional experiences. These emotions are known as epistemic emotions (Pekrun & Stephens, 2012; Pekrun, Vogl, et al., 2017). The object focus of epistemic emotions is knowledge and knowledge generation (Brun et al., 2008), which can trigger different discrete emotions, such as surprise, curiosity, and confusion (Vogl et al., 2020; Vogl et al., 2019). Epistemic emotions also have been linked to cognitive performance. For instance, research has found that epistemic emotions predicted processes in self-regulated learning, which in turn predicted complex mathematical problem solving (Muis, Psaradellis, et al., 2015) and learning outcomes on climate change (Muis, Pekrun, et al., 2015).

Further, working on such cognitive tasks has been found to arouse subjective experiences of high mental effort and corresponding mental fatigue during experimental task execution (Gergelyfi et al., 2015; Hopstaken et al., 2015; Hopstaken et al., 2016). These mental states are not classically considered emotional states like anger or pride, but they in fact are closely conceptually related in that they may be experienced as aversive and have specific physiological components.

It becomes apparent that tasks can trigger emotions in people, which in turn are linked with how they perform on these tasks. However, to date, these processes have – to our knowledge – not been examined, nor considered, in laboratory settings designed to measure VWM functions.

**The Present Study**

We report the results of four consecutive studies. In Study 1, we sought to explore qualitatively what participants feel when performing a VWM task (Research Question (RQ) 1a). In doing so, our goal was to discern discrete emotions that participants may experience, such as joy, pride, anger, and boredom. Further, we sought to explore which aspects of the VWM task triggered these discrete emotions (RQ 1b).

Studies 2 and 3 used a quantitative design to explore, whether and how task-induced emotions are linked to VWM performance, both on an inter-individual level (Study 2; RQ 2a), and on an intra-individual level (Study 3; RQ 2b). Additionally, in Study 3, we explored whether a typical experimental manipulation (set-size) had any systematic effects on the emotions experienced during the task. Study 4 was a high-powered replication ($N$ = 110) of Study 3 that also exploratorily addressed a few interesting subsidiary questions that came up during the review process.

The research reported herein was conducted in accordance with the ethical standards expressed in the Declaration of Helsinki and has received a formal waiver of ethical approval by the ethics committee of the Department of Psychology, LMU München. Participation was voluntary and written informed consent was obtained from all participants for each study. The data files can be found on the Open Science Framework (Laybourn et al., 2020).

**Study 1**

To explore which discrete emotions are experienced when engaging in a VWM task, participants in this study were asked to verbalize their feelings and any related thoughts thereof while performing the continuous color-report task (Zhang & Luck, 2008). This approach is based on the think aloud method by van Someren et al. (1995), which is traditionally used when trying to identify and understand underlying processes in problem solving by encouraging participants to verbalize their thoughts and strategies while trying to solve a certain problem. We deemed it suitable also for identifying the emotional experiences and related thoughts thereof participants encountered when performing a VWM task (see also e.g., D'Mello and Graesser, 2012, who refer to this as "emote-aloud").

We chose the continuous color-report task as it is a well-established paradigm used in VWM research and has consistently been reported as a highly reliable method for estimating individual differences in VWM. The task involves memorizing multiple, shortly presented visual stimuli (typically two or more colored squares, the sample array), and then, after a short retention period of around 1s, being prompted to recall one of them (the test array; here, a thick black frame indicates which of the squares from the sample array should be recalled) by selecting the color of the prompted square on a continuous color wheel which surrounds the test array (see Figure 1).

**Methods**

*Sample*

Participants of this study were $N$ = 19 (11 female; $M_{age}$ = 30.21; $SD$ = 8.49), who all stated not to be color-blind and to have normal or corrected-to-normal vision.

*Continuous Color-Report Task Specifications*

All stimuli were generated in MATLAB using the *Psychophysics Toolbox*. Stimuli were presented on 24" TFT-LCD monitors (ASUS VG248QE, 1920×1080 pixels, 60 Hz) at a viewing distance of approximately 70 cm. The to-be-remembered stimuli were five colored squares (1 x 1°) on a dark grey background (RGB: 60, 60, 60), which randomly (with a distance of at least 1.5° between each) appeared at

**Figure 1**

*Continuous Color-Report Task Showing Set Size 4*
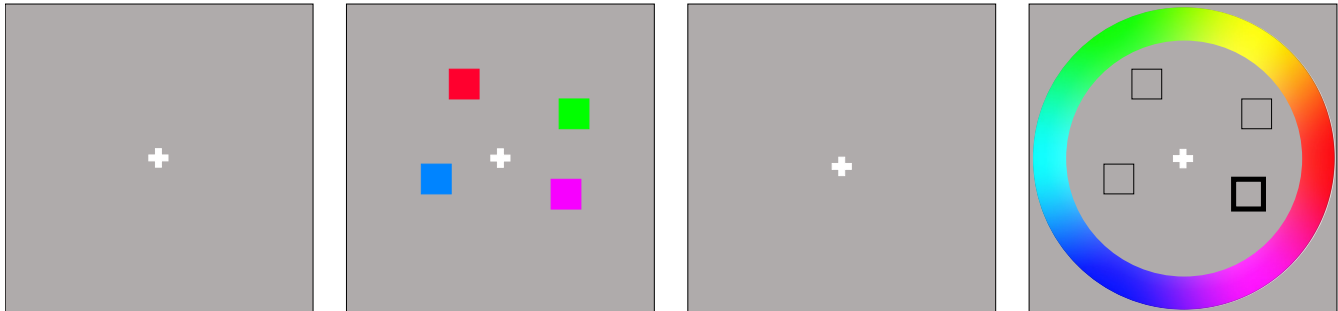


least 2° from the white fixation cross (0.5°) within a rectangular region with a width and height of $10 \times 9°$, centered on the fixation cross. Colors were randomly drawn from a circle with a radius of 40 in a luminance plane of the CIE 1976 L*a*b* color space (L* = 63, center: a* = 9, b* = 27, illuminant: D65, 2° standard observer).

All trials followed the same order: After the fixation cross was briefly presented for 1000 ms (inter-trial interval), the sample array appeared for 1000 ms containing five colored squares. After a 1000 ms delay interval, participants were presented with the test array, which remained until a response was given. Responses were not timed. In keeping with the traditional paradigm, performance feedback was not provided.

### Procedure and Think Aloud Protocol Transcription

Participants performed the task alone in the laboratory with only the researcher present. First, participants were given the opportunity to familiarize themselves with the VWM task by reading a short introduction to the task and performing a practice block containing 20 trials. The actual VWM task consisted of two blocks, each containing 165 trials.

The first block served the purpose of the participants' full habituation of the tasks' procedure before asking them to emote aloud during the task, thus in this block, participants did the task in silence. During the second block, the emote aloud procedure was introduced. To this end, participants were prompted with the following instructions: "You will now complete another block of trials, just like the one you have just completed. During this next block, we ask you to please verbalize any feelings or thoughts you are experiencing during the task". The resulting monologue was recorded and later transcribed verbatim. If required, the researcher reminded participants to "think and feel aloud". Overall, participants took approximately between 45 and 90 minutes to complete the task.

### Analysis and Coding Reliability

Data analysis was carried out on the basis of qualitative content analysis proposed by Mayring (2014) using the open access web-application QCAmap (Mayring & Fenzl, 2014).

As to the authors' knowledge, there are no prior findings regarding participants' affective states when performing a VWM task in the laboratory, categories needed to be extracted from the textual material itself, using the content analytical technique of inductive category formation (Mayring, 2014).

To determine what participants felt when performing a VWM task (RQ 1a), we categorized all text passages in which "participants referred to emotional states experienced during the VWM task" (selection criterion). Next, categories were phrased as "specific emotional states or personal feelings, which participants referred to during the VWM task" (level of abstraction).

To explore which aspects of the VWM task triggered these discrete emotions (RQ1b), the selection criterion was set to text passages in which "participants explicitly referred to or hinted at reasons, sources or processes related to the emotional states or feelings experienced/perceived during the VWM task". Categories were phrased as "specific reasons, sources or processes affecting or leading to participants' emotional states and feelings experienced/perceived during the VWM task" (level of abstraction). Further details of the coding procedure can be found in the Supplemental Material.

In order to establish the degree of reliability for the category system and reproducibility of the categorizations, a subsample of six randomly chosen interviews (approx. 33% of the entire material) were coded by two independent coders. The coders reached substantial agreement for both research questions (RQ1a: $K = .79$ (95% *CI*, .68 to .89), $p > .01$; RQ1b: $K = .72$ (95% *CI*, .61 to .82), $p > .01$).

**Table 1**

*Discrete Emotions Stated by Participants During a VWM Task in Order of Frequency*

| Discrete emotion | Frequency by participants | Absolute count |
|---|---|---|
| Anger | 8 participants | 24 |
| Frustration | 7 participants | 20 |
| Joy | 7 participants | 17 |
| Boredom | 6 participants | 9 |
| Tension/Nervousness | 3 participants | 3 |
| Confusion | 2 participants | 2 |
| Desperation | 2 participants | 5 |
| Hope | 2 participants | 3 |
| Same | 1 participant | 2 |
| Disappointment | 1 participant | 1 |
| Uncertainty | 1 participant | 1 |
| Anxiety | 1 participant | 1 |

## Results and Discussion

Overall, 12 discrete emotions were identified (see Table 1). The most frequently stated emotions were anger (stated by eight participants), frustration (seven participants), joy (seven participants), and boredom (six participants). Participants varied in terms of how many different emotions they reported during the task (ranging from one to six emotions) and how often they reported to experience them (ranging from one to nine times).

These findings confirm previous speculations and provide first empirical evidence that participants experience emotions during a VWM task, which occur because of the task itself. This is in line with emotion research proposing that emotions can be triggered by (cognitive characteristics of) a task (Pekrun, Vogl, et al., 2017).

In regards to RQ1b, 10 categories were identified, which involved underlying processes associated with participants' emotional experiences during the VWM task (see Table 2). Participants' *self-expectations regarding the VWM task* were identified most frequently to be related to their emotional experiences (eleven participants, referred to 28 times). The majority of the coded passages in this category indicated that participants wanted and also expected to do well on the task, as it was perceived initially as a simple task. However, the task proved to be more difficult than expected. This led to participants experiencing negative emotions, such as anger ("You get angry when you don't know it because this isn't really that difficult, actually", Participant D).

Eleven participants perceived the *VWM task as challenging in a negative sense*. Here, participants reported the VWM task as being stressful or overwhelming for them and resulting in negative emotions for the individual, as is illustrated by the following examples: "I always try to remember, more

or less, the general color. And when each one is different, then I am out of my depth.... That just makes me angry" (Participant C).

"But sometimes, I don't know, sometimes the time to look at the colors is too short and then I get desperate because I try to recite the colors and to see which ones come in pairs and when there are many different colors, all of a sudden nothing works anymore." (Participant G)

Six participants referred to being *dissatisfied with the VWM task design*, which mainly resulted in frustration: "What's also frustrating is the cross in the middle of the screen" (Participant A). One participant reported to experience boredom: "Because this is the second round, you just start noticing that it is starting to be boring, yes, because it is always the same thing" (Participant K).

Participants' *general judgement of the VWM task* (six participants) pertained to the general attitude they reported having towards the task. For instance, some participants referred to the task as being "pointless" or "silly" (Participant K), others compared the task to a game (Participant M) or an exam situation (Participant P). A *change in motivation* was also referred to by six participants, especially towards the end of the VWM task: "For some reason I'm, I'm starting to notice that it doesn't matter to me that much anymore, I am not clicking anywhere specific anymore" (Participant S).

Pertaining to the category *social comparison*, six participants wondered how they were performing in comparison to the other participants: "I'm always asking myself, if I am that bad or if the others are also this bad." (Participant P). One participant stated to be angry for comparing their own achievement to those of others.

Four participants referred to *missing performance feedback*, which was mainly associated with interest: "It would be interesting to know your score, at the end. Maybe, I don't know, a smiley face indicating whether you were right or wrong" (Participant D).

Four participants also referred to a *missing time reference*, which seemed to be frustrating, as the following example illustrates: "You don't know when it'll be over. I think that's what's bugging me" (Participant E).

Three participants perceived the *VWM task as challenging in a positive sense*. Here, participants mainly reported that the VWM task fueled their ambition to perform well, but when they did not, they experienced negative emotions, for example:

"I am still ambitious. It's not as if I would stop doing this straight away. I want to continue doing this and I want to do well at this and I try every time again and again. But somehow you still are disappointed when you don't know the answer." (Participant D)

Finally, three participants referred to *strategies to improve their achievement*: "You start and build themes and then you wait and see, and then you try and do it well" (Participant E).

**Table 2**

*Underlying Processes Regarding Emotional Experiences in Order of Frequency*

| Underlying process | Frequency by participants | Absolute count |
| --- | --- | --- |
| Self-expectations | 11 participants | 28 |
| VWM task is challenging (negative sense) | 11 participants | 23 |
| General judgment of the VWM task | 6 participants | 27 |
| Dissatisfaction with the VWM task design | 6 participants | 12 |
| Change in motivation | 6 participants | 11 |
| Social comparison | 6 participants | 10 |
| Referring to missing performance feedback | 4 participants | 9 |
| Referring to missing time reference | 4 participants | 7 |
| VWM task is challenging (positive sense) | 3 participants | 8 |
| Thoughts on strategies to improve achievement | 3 participants | 3 |

It is worth noting that the think and emote aloud method relied upon participants to be able to register what they are thinking and feeling on a meta-cognitive level and to translate these complex internal processes into words for a third party to understand. A few participants seemed to struggle with this in that they did not verbalize any emotional experiences, others described the think aloud task itself as a source for certain emotional experiences. Yet we took great care not to categorize emotional experiences which were triggered by the think aloud method.

Overall, the results showed that participants experienced an array of different discrete emotions while performing the continuous color-report task, which varied between individuals – some seemed to enjoy the task more, yet others found it frustrating. It also became evident that participating in the continuous color-report task implied going through highly intra-individually varying emotional states – at some points during the experimental block, participants were activated, engaged, and enjoying it, while next they had trouble focusing, worried about their achievement, and became frustrated.

Only few statements pertained to epistemic emotions, such as confusion (Vogl et al., 2020), and some emotions seemed to have been triggered by certain experimental design features we had realized (no performance feedback, no explicit breaks). Most importantly, though, it became evident that the dominating theme of most participants' thinking while engaging in the continuous color-report task was subjective success and failure. This confirms speculations expressed earlier for example by Rouder and colleagues (2008) that some participants can be "intimidated" by the task, and Spachtholz et al. (2014) who conjectured that the VWM task can "signal requirements" which trigger emotions in the participants. Clearly, the key task requirement built into a memory task such as the continuous color-report task is to re-member the "correct" color, so participants are fully aware that they can fail versus succeed at each trial. As such, a key insight from this qualitative study was that this task clearly places participants into an achievement situation. That is, participants find themselves in a situation where judgments regarding achieving or failing against some standard, be it task-based, self-based, or other based (Elliot et al., 2011) are dominant.

## Study 2

Previous research has demonstrated that emotions induced prior to a VWM task can affect VWM performance. In a second study, we sought to establish whether emotions, which were induced by the VWM task itself, were also systematically linked to VWM performance. To this end, participants performed the continuous color-report task and were asked to rate their emotional experiences during the task immediately after. We opted for such a summative, retrospective task emotion assessment in order to minimize any disruptions during the task.

Taking up the results of Study 1, we chose to assess joy, anger, frustration, and boredom. As joy was the only positive emotion explicitly labeled by the participants in Study 1, and we sought to assess diverse discrete emotions also of positive valence (Pekrun, 2018; Posner et al., 2005), we chose to additionally include pride in this study. Pride is an important self-conscious emotion and a prototypical achievement emotion tightly linked with appraisals of task success (Lagattuta & Thompson, 2007), and given the situational salience of achievement we had identified in Study 1, we deemed it promising to further investigate the link between this emotion and VWM performance.

## Method

### *Sample*

Forty-seven individuals initially participated in this study (31 female; $M_{age}$ = 26.09; $SD$ = 3.85) who all stated not to be color-blind and to have normal or corrected-to-normal vision. They received either course credit or monetary compensation for their time.

Two individuals were excluded from further analysis due to substantial missing data in the self-reported emotions and technical difficulties during the VWM task. No participants were excluded due to extremely poor performance. The final sample of this study thus was $N$ = 45 (30 female; $M_{age}$ = 26.24; $SD$ = 3.86).

### *Procedure, Stimuli and Measurements*

**Procedure.** Participants familiarized themselves with the VWM task by performing a short practice block containing 30 trials, which started either with 15 trials displaying four squares followed by 15 trials displaying eight squares, or vice versa. The actual task contained 240 trials, which were arranged in alternating blocks of 30 trials each, displaying either four- or eight-square arrays (set-size; starting size counterbalanced). At the end of the VWM task, participants were prompted to fill in a paper-and-pencil questionnaire asking them to report, retrospectively, how they felt during the VWM task.

**Continuous Color-Report Task.** We employed the same task as in Study 1, except that we deliberately introduced variability in the difficulty of the task by using two different sample array set-sizes: four or eight colored squares. We did so because on the one hand, we wanted participants to be fairly comfortable with the task during certain phases of the experiment (set-size 4). On the other hand, we wanted to place them systematically in demanding achievement situations (set-size 8; e.g., Rouder et al., 2008), as results from Study 1 indicated that emotions participants experienced during the VWM task were linked to its achievement requirements. Further, it is common to vary set-size in paradigms designed to measure VWM. No performance feedback was provided and responses were not timed. Participation length for the 240 trials ranged between 25 and 35 minutes, approximately.

VWM performance was operationalized by computing the absolute angular distance (in degrees) between the probed item's color (on the color wheel) and the selected color (henceforth: recall error) for each participant.

**Emotion Ratings via Paper-and-Pencil Questionnaire.** To assess participants' emotions regarding the VWM task, we asked participants to rate five items concerning the emotions they experienced during the VWM task (*I enjoyed the task*, *I felt proud*, *I felt angry*, *I was frustrated*, *I felt bored*) on a five-point Likert type scale ranging from strongly disagree to strongly agree. These judgments were made immediately after finishing the experiment.

## Results and Discussion

As expected, there were no significant effects of set-size order on either recall error or self-reported emotions (all $t$s < .81, all $p$s > .37) and we therefore do not consider this group factor further. An overview of the descriptive statistics for the emotion measures in Study 2 can be found in Table 3.

As the emotions were rated with single items on a 5-point Likert scale, thus not affording an interval-scale measurement level, we obtained Spearman's correlation coefficient for VWM performance and the emotions joy, pride, anger, frustration, and boredom to explore the link between task-induced emotions and VWM performance (recall error) on an inter-individual level (RQ2a; see Table 4). Results indicated that there was a significant negative association between recall error and joy ($r_s$ = −.34, 95% BCa *CI* [−.606, −.017], $p$ = .02) and pride ($r_s$ = −.30, 95% BCa *CI* [−.602, .015], $p$ = .04). Further, there were significant positive associations between recall error and anger ($r_s$ = .32, 95% BCa *CI* [.026, .561], $p$ = .03), frustration ($r_s$ = .34, 95% BCa *CI* [.048, .555], $p$ = .03), and boredom ($r_s$ = .33, 95% BCa *CI* [.017, .57], $p$ = .03)[1].

The results imply that the better participants performed on the VWM task relative to others, the more joy and pride they experienced. However, when participants performed more poorly on the task relative to others, they experienced more anger, frustration and boredom. Importantly, though, these are purely correlative findings, so we hasten to caution the reader (and ourselves) not to interpret these findings as causal relations; emotions might affect performance, performance might affect emotions or there might be a third variable that affects both emotions and performance. Even more likely, the correlation might be due to a complex reciprocal interaction between performance, emotions, and maybe additional variables (as discussed in more detail in the Introduction and General Discussion sections with regard to Pekrun's (2006) control-value theory).

## Study 3

The results from Study 1 had suggested that not only did the participants differ from one another, that is, on an inter-individual level, with respect to their task-induced emotions during the VWM task (as followed up upon in Study 2), but also, single individuals seemed to experience widely ranging levels of emotions during the task. Additionally, Study 1 had revealed that participants were mostly preoccupied with their subjective performance during the task, and the emotions they experienced could be largely classified as achievement emotions. Following up on this, in Study 3, we sought to explore whether task difficulty (i.e., set-size) had an effect on participants' emotions, and we aimed to assess the intra-individual variation of emotional experiences during the continuous color-report task execution. Further, we aimed to explore whether and how these varying emotions related to VWM performance, on

---

[1]We confirmed that all those effects were significant ($p$ < .05) even when adopting the Benjamini-Hochberg procedure for ruling out false discoveries.

**Table 3**

*Descriptive Statistics for Emotional Measures in Studies 2, 3, and 4*

| | Study 2 (N = 45) | | Study 3 (N = 44) | | | | Study 4 (N = 110) | | | | | |
| | Single retrospective measure | | Mean score across the blockwise measures | | | | Single retrospective measure | | Mean score across the blockwise measures | | | |
| | Median | Range | M (SD) | Range | Skew. | Kurt. | Median | Range | M (SD) | Range | Skew. | Kurt. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Joy | 3.0 | 1–5 | 2.5 (0.73) | 1–4 | −0.17 | −0.35 | 3.0 | 1–5 | 2.62 (1.02) | 1–5 | 0.49 | −0.34 |
| Pride | 2.0 | 1–5 | 2.1 (0.67) | 1–3.5 | 0.07 | −0.85 | 2.0 | 1–5 | 2.22 (0.77) | 1–5 | 1.08 | 2.01 |
| Anger | 2.0 | 1–4 | 1.9 (0.90) | 1–4.3 | 0.71 | −0.48 | 2.0 | 1–5 | 2.08 (0.98) | 1–4.6 | 0.66 | −0.32 |
| Frustration | 3.0 | 1–5 | 2.4 (0.93) | 1–4.9 | 0.47 | 0.07 | 3.0 | 1–5 | 2.76 (1.20) | 1–5 | 0.19 | −0.70 |
| Boredom | 3.0 | 1–5 | 2.9 (0.88) | 1.1–4.8 | 0.17 | −0.51 | 3.0 | 1–5 | 3.02 (1.17) | 1–5 | −0.10 | −0.98 |

an intra-individual level. To this end, in Study 3, we asked participants to rate their emotional experiences at multiple time points during the VWM task. In other words, we sought to explore whether any dynamics of the participants' emotions across the course of the experiment, as assessed through multiple emotion ratings after short experimental sub-blocks, co-fluctuated with the dynamics of the performance across those sub-blocks, within the participants.

## Method

### Sample

Forty-six individuals (26 female; $M_{age}$ = 25.57; $SD$ = 3.86) initially participated in this study, who all stated not to be color-blind and to have normal or corrected-to-normal vision. They received either course credit or monetary compensation for their time.

One person was excluded from further analysis due to substantial missing data in the self-reported emotions and one participant was excluded due to extremely poor performance (average recall error of more than two standard deviations above the mean), indicating poor study commitment and thus low overall data quality. The final sample of the study thus was $N$ = 44 (24 female; $M_{age}$ = 25.50; $SD$ = 3.89).

### Procedure, Stimuli, and Measurements

**Procedure.** We largely adopted the same procedure as described in Study 2. The key difference was that in this study, participants were prompted to rate the emotion items presented to them in a paper-and-pencil questionnaire several times within the experiment (after each block of 30 trials displaying either four- or eight-square arrays), as we sought to assess the potential intra-individual variability in emotions regarding the VWM task execution in the course of the experiment.

**Continuous Color-Report Task.** We adopted the same continuous color-report task as described in Study 2, and operationalized VWM performance in terms of recall error accordingly. We thus obtained recall error for each block of 30 trials.

**Emotion Ratings via Paper-and-Pencil Questionnaire.** To assess participants' emotions regarding the VWM task on an intra-individual level, we asked participants "*How are you currently feeling?*" at eight time points during the VWM task. At each time point participants were asked to rate five items concerning the emotions they experienced during the VWM task (*I am enjoying the task*, *I feel proud*, *I feel angry*, *I am frustrated*, *I feel bored*) on a five-point Likert type scale ranging from strongly disagree to strongly agree.

### Analysis Approach

We calculated random intercept, fixed-slope models for each emotion/recall error combination. For this analysis, all emotion and recall error scores were transformed into $z$-scores, so that the within-person regression parameters could be interpreted as standardized correlations, to allow for comparability between the results from this study and Study 2.

**Table 4**

*Spearman Correlations from Study 2 and Intra-Personal Correlations from Study 3 and 4 between Emotions and Recall Error*

| | Study 2 | | Study 3 | | | Study 3 (controlling for set size) | | | Study 4 | | | Study 4 (controlling for set size) | | | Study 4 (controlling for set size and pre-experimental mood) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $r_s$ | 95% CI | $r$ | $SE_\beta$ | 95% CI | $r$ | $SE_\beta$ | 95% CI | $r$ | $SE_\beta$ | 95% CI | $r$ | $SE_\beta$ | 95% CI | $r$ | $SE_\beta$ | 95% CI |
| Joy | −.34* | −.61, −.02 | −.31** | .05 | −.42, −.21 | −.11** | .04 | −.18, −.04 | −.29** | .04 | −.37, −.21 | −.11** | .03 | −.18, −.05 | −.11** | .03 | −.18, −.05 |
| Pride | −.30* | −.60, .02 | −.42** | .05 | −.53, −.32 | −.16** | .04 | −.23, −.08 | −.38** | .04 | −.45, −.30 | −.19** | .03 | −.25, −.14 | −.20** | .03 | −.25, −.14 |
| Anger | .32* | .03, .56 | .25** | .06 | .12, .37 | .07 | .04 | .01, .15 | .26** | .04 | .18, .34 | .10** | .03 | .04, .16 | .10** | .03 | .05, .16 |
| Frustration | .34* | .05, .55 | .38** | .06 | .26, .50 | .16** | .04 | .08, .24 | .33** | .04 | .25, .40 | .13** | .03 | .07, .19 | .13** | .03 | .07, .19 |
| Boredom | .33* | .02, .57 | .21** | .06 | .10, .32 | .13** | .04 | .06, .20 | .15** | .04 | .07, .23 | .05 | .03 | −.006, .11 | .05 | .03 | −.007, .11 |

*Note.* Intraindividual correlations in Study 3 and 4 were obtained via two-level hierarchical models.

*p < .05 **p < .01

## Results and Discussion

Descriptive statistics of all emotion measures in Study 3 are shown in Table 3. As in Study 2, there were no significant between-person effects of set-size order on either recall error or self-reported emotions (all $ts < 1.52$, all $ps > .14$). Further, to test whether the multiple emotion ratings affected VWM performance, an independent-samples $t$-test was calculated using participants' average recall error from Study 2 ($M = 47.02$, $SD = 8.54$) and Study 3 ($M = 47.75$, $SD = 10.89$). Results yielded no significant difference between the two sample means ($t(87) = -0.35$, $p = .73$, $d = -0.07$).

Next, we explored effects of set-size. As could be expected, paired-sample t-tests showed significant effects of set-size on recall error ($t(43) = -24.17$, $p < .01$, $d_z = -3.64$) thus, easier blocks (i.e., arrays with four squares) were associated with better performance ($M_4 = 35.70$, $SD_4 = 10.49$; $M_8 = 59.79$, $SD_8 = 12.22$). Regarding the emotions, paired-samples $t$-tests further showed significant effects of set-size on joy ($t(43) = 4.66$, $p < .01$, $d_z = 0.70$), pride ($t(43) = 5.96$, $p < .01$, $d_z = 0.90$), anger ($t(43) = -3.91$, $p < .01$, $d_z = -0.59$) and frustration ($t(43) = -5.00$, $p < .01$, $d_z = -0.75$), indicating that the variation in task difficulty affected how participants felt while performing the VWM task. Specifically, when confronted with set-size 4, participants experienced more enjoyment ($M_4 = 2.70$, $SD_4 = 0.82$) and pride ($M_4 = 2.32$, $SD_4 = 0.79$) than when confronted with set-size 8 (joy: $M_8 = 2.31$, $SD_8 = 0.74$; pride: $M_8 = 1.85$, $SD_8 = 0.65$). Further, participants experienced less anger ($M_4 = 1.72$, $SD_4 = 0.82$) and frustration ($M_4 = 2.15$, $SD_4 = 0.91$) when confronted with set-size 4 than with set-size 8 (anger: $M_8 = 2.06$, $SD_8 = 1.02$; frustration: $M_8 = 2.60$, $SD_8 = 1.04$). The effect of set-size on boredom did not reach significance ($t(43) = -1.97$, $p = .06$, $d_z = -0.30$; $M_4 = 2.88$, $SD_4 = 0.95$; $M_8 = 3.06$, $SD_8 = 0.91$; see Figure 2).

Regarding the relationship between performance and emotions, results showed that every discrete emotion significantly correlated with recall error on the intra-individual level (see Table 4). Specifically, joy and pride were significantly negatively linked with recall error, indicating that when a participant performed better on a block of 30 trials than on the other blocks, they enjoyed it more, and were more proud, relative to the other blocks. Vice versa, when participants experienced more of those positive emotions during a block, they performed better on the task on that particular block relative to the other blocks. In turn, anger, frustration, and boredom were significantly positively linked with recall error, implying that the worse a participant performed on a particular block the more they experienced those negative emotions during that block, and vice versa. Finally, taking into consideration that set-size also had a strong effect on the emotions, in a last analysis step, we additionally considered set-size in our analyses. This implied exploring whether, even in blocks of the same difficulty, participants' emotions varied, and whether this variation in task experiences was systematically related to performance. Intriguingly, the effects remained significant for all discrete emotions except for anger (see Table 4). This implies that anger was strongly driven by task difficulty. However, participants' fluctuations in joy and pride across the experiment were still significantly negatively related with their performance in the task, and their fluctuations in frustration and boredom were positively related with performance above and beyond the performance variability induced by the array size manipulation.

**Figure 2**

*Effects of Set Size on Emotions in Study 3*
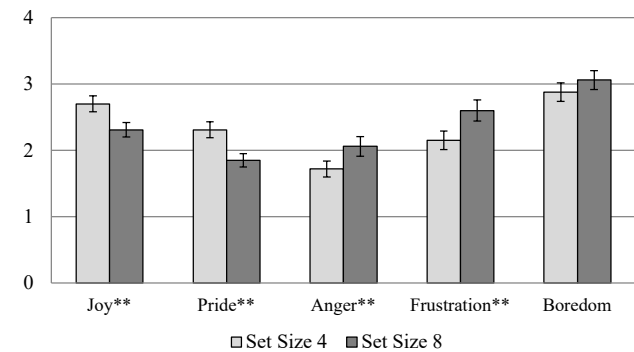


*Note.* \*\*$p < .01$; error bars display ± one standard error.

## Study 4

This study was designed as a high-powered replication of Study 3 (intra-individual links between emotions and performance). Due to the COVID-19 pandemic, data collection in the laboratory was not possible. The study was therefore moved from a laboratory to an online environment and data was collected via the data collection platform Prolific (thus drawing from a different population). Otherwise we aimed to replicate Study 3 as closely as possible. The functionality of the online environment for the assessment of both the continuous color-report paradigm as well as self-report items of emotional experiences after each 30-trial sub-block was piloted.

In addition to the emotion ratings after each block, we also included further "overall retrospective" emotion rating items at the very end of the experimental session. We were unsure whether participants were able to sufficiently abstract from their current emotional state when answering these items after having been probed for their current emotional state multiple times during the experiment. If they were able to do so, this would provide a relatively cost-free opportunity to additionally replicate Study 2 conceptually, because the additional rating would not affect the main data of interest (related to the replication of Study 3).

As a further extension of this replication, we had participants fill out the Positive and Negative Affect Schedule (PANAS; Watson et al., 1988) before the start of the experiment. This would allow us to take their pre-experimental mood into account as well. As such, we could explore whether mood before the experiment had any effects on performance, and whether the task-induced emotion—performance links as demonstrated in Study 3 would persist above and beyond any such effects of mood.

### Method

#### Sample

We used the effect of set-size on boredom in Study 3 ($d_z = -0.30$), which had just failed to reach significance, to determine the necessary sample size for this replication. Using *G\*Power* aiming

at power $1 - \beta = .9$; $\alpha = .05$ for this effect, we obtained a required $N$ of 97. Administering the study online via Prolific involved making batches of experimental slots available at several standardized time slots, which would not necessarily always fill up. We further anticipated that we could not include all initial participants in our analyses due to exclusion criteria, potential technical problems in the online data collection environment, or other factors and therefore slightly oversampled. Our final initial sample (before exclusion) was $N = 124$ participants (43 females, 1 other; $M_{age} = 26.47$; $SD = 7.76$).

We took advantage of Prolific's option to preclude participants from study-participation due to certain factors. Therefore, all participants reported not to be color-blind and to have normal or corrected-to-normal vision. One participant was excluded from the study due to technical problems, which lead to substantial missing data. We applied the same exclusion rule for sorting out participants based on extremely low performance as in Study 3 (recall error of more than two standard deviations above the mean), indicating poor study commitment and thus low overall data quality. On that basis, four participants were excluded from further analyses. Further nine participants were excluded for taking too long on the experiment overall (more than two standard deviations above the mean experiment duration, i.e., more than 86 minutes), an exclusion rule we deemed necessary given the highly uncontrolled digital setting. The final sample included for analysis was thus $N = 110$ participants (38 female; $M_{age} = 26.56$; $SD = 7.92$), thus slightly oversampling relative to the required N as implied by the power analysis for the reasons stated above.

### Procedure, Stimuli, and Measurements

**Procedure.** Data collection occurred online via Prolific. First, participants were prompted to rate items on the Positive and Negative Affect Schedule (PANAS; Watson et al., 1988) before they were introduced to the same continuous color-report task used in studies 2 and 3. Identical to the procedure in Study 3, after a short practice block of 30 trials (15 of each set-size), participants performed the experimental block of 240 trials consisting of 30 trials each in a row with array size 4 or 8. As results from both Study 2 and 3 yielded no significant effects of set-size starting order on either recall error or self-reported emotions, this was not counterbalanced in the present study and all participants started with set-size 4. As in Study 3, after each of these sub-blocks of 30 trials, participants were prompted to rate the emotion items; this time presented to them on the screen in the online environment. Different from Study 3, attempting to potentially also replicate our findings from Study 2, at the very end of the block of 240 trials, participants were prompted to report, retrospectively, how they felt during the entire experiment. Participants received monetary compensation for their time. The various measures are described in more detail in the following.

**Continuous Color-Report Task.** We adopted the same continuous color-report task as described in Studies 2 and 3. VWM performance was operationalized in terms of recall error accordingly. We obtained the average recall error for each block of 30 trials (for the intra-individual analyses) and the overall average recall error after 240 trials (for the inter-individual analyses).

**Positive and Negative Affect Schedule.** To assess participants' mood prior to the VWM task, the Positive and Negative Af-

fect Schedule (PANAS; Watson et al., 1988) was used. The PANAS is suitable to either measure trait affectivity (personal tendencies to experience generally positive vs. negative affect) or state affect (i.e., momentary positive versus negative affect/mood). It includes 20 items, ten of which are positive (e.g. *interested*, *inspired*, *active*) and ten are negative (e.g., *distressed*, *hostile*, *jittery*). For the present purpose, we used the state instructions, asking participants to "Indicate to what extent you feel this way right now, that is, at the present moment." Responses were given on a five-point Likert type scale ranging from *very slightly or not at all* to *extremely*.

**Emotion Ratings during the VWM task.** Exactly as in Study 3, to assess participants' emotions regarding the VWM task on an intra-individual level, we asked participants, "*How are you currently feeling?*" after each sub-block (eight times in total). At each time point participants were asked to rate five items concerning the emotions they experienced during the VWM task (*I am enjoying the task*, *I feel proud*, *I feel angry*, *I am frustrated*, *I feel bored*) on a five-point Likert type scale ranging from *strongly disagree* to *strongly agree*.

**Overall Emotion Ratings of the VWM task.** Exactly as in Study 2, we asked participants to retrospectively rate their emotions after they completed the entire VWM task by asking, "When looking back across the entire memory task: how were you feeling during the task?" Participants again rated five items (*I enjoyed the task*, *I felt proud*, *I felt angry*, *I was frustrated*, *I felt bored*) on a five-point Likert type scale ranging from *strongly disagree* to *strongly agree*.

### Results and Discussion

To explore whether the online setting had any effects, we tested average recall error means from Study 3 against recall error means from Study 4. Participants of this online study performed slightly better overall ($M = 43.25$, $SD = 12.28$) than participants of the comparable laboratory study (Study 3, $M = 47.75$, $SD = 10.89$), $t(152) = 2.12$, $p = .04$, $d = 0.38$.).

An overview of the descriptive statistics for the emotion measures in Study 4 can be found in Table 3. Somewhat disappointingly, but not unexpectedly (see above), the Spearman's correlation coefficients for overall VWM performance and the overall retrospective emotion judgements for joy, pride, anger, frustration, and boredom ($r_s = -.06/-.11/.16/.04/.05$; $p = .53/.24/.11/.66/.61$; respectively) did not reach statistical significance. Thus, simply adding an extra question at the end of the experiment did not serve the purpose of replicating Study 2. This is likely due to a crucial change in the study design (which was introduced because our main aim was to replicate Study 3): As opposed to Study 2, where participants performed the VWM task without any interruption and were asked to rate one item regarding their achievement emotions towards the task in one single retrospect across the entire experimental block, participants in Study 4 were asked to continually rate their current emotional experiences eight times within the VWM task, before rating the overall emotion item retrospectively. It appears that the multiple question rating interfered with the final retrospective question. Maybe this made the variation of set-size and corresponding emotional variation particularly salient to the participants, or they were unable to sufficiently abstract from their current emotional state after getting used to report exactly this.

In order to replicate the findings from Study 3, we again calculated random intercept, fixed-slope models for each emotion/recall error combination, using the emotion and recall error z-scores, to allow for comparability. Results showed that, as in Study 3, every discrete emotion significantly correlated with recall error on the intra-individual level (see Table 4). For joy, pride, anger, and frustration this also held true when controlling for set-size. For boredom, the correlation with recall error was no longer significant when controlling for set size. Additionally, in a third step of our multilevel regression analyses, we also considered mood before the experiment on the inter-person level (level 2), which did not alter the results (see Table 4). In sum, results of Study 3 were fully replicated for enjoyment and pride (positive intra-individual links with performance) and frustration (negative intra-individual links with performance). Intriguingly, in contrast to Study 3, the negative anger—performance link now remained significant when controlling for set-size (as well as mood), but the boredom—performance link proved no longer significant with this control in this high-powered study.

Paired-samples t-tests also replicated the findings from Study 3 that set-size had a significant effect on each of the emotions (joy: $t(109) = 8.10, p < .01, d_z = 0.77; M_4 = 2.78, SD_4 = 1.04, M_8 = 2.45, SD_8 = 1.05$; pride: $t(109) = 8.54, p < .01, d_z = 0.81; M_4 = 2.40, SD_4 = 0.79, M_8 = 2.05, SD_8 = 0.81$; anger: $t(109) = -6.22, p < .01, d_z = -0.59; M_4 = 1.94, SD_4 = 0.91, M_8 = 2.23, SD_8 = 1.10$); frustration: $t(109) = -7.27, p < .01, d_z = -0.69; M_4 = 2.55, SD_4 = 1.12, M_8 = 2.97, SD_8 = 1.16$; including boredom: $t(109) = -5.14, p < .01, d_z = -0.48, M_4 = 2.92, SD_4 = 1.16, M_8 = 3.12, SD_8 = 1.21$). Thus, the non-significance of an effect of set-size on boredom in Study 3 (on which we had based the power analysis to determine the sample size in Study 4) was indeed simply a power issue that was resolved in Study 4. These results indicate once more that the variation in task difficulty affected how participants felt while performing the VWM task (see Figure 3).

## Figure 3

*Effects of Set Size on Emotions in Study 4*
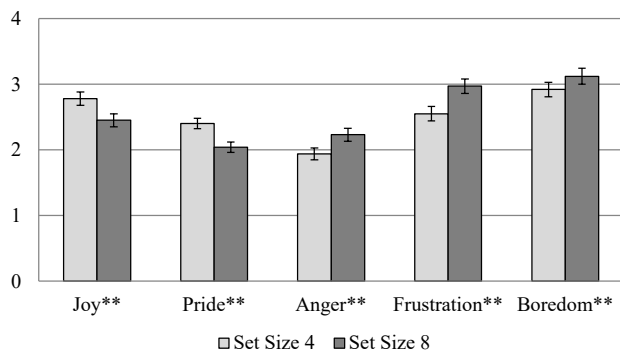


*Note.* **$p < .01$; error bars display $\pm$ one standard error.

## General Discussion

The present research aimed to explore whether participants experience emotions during a VWM task, which are induced by the task itself, and which discrete emotions these are. Next, and most importantly, we sought to demonstrate that such incidentally emotions were systematically linked with VWM performance (positive links for pleasant emotions, negative links for unpleasant emotions). To this end, we conducted an exploratory qualitative study, followed up by three confirmatory quantitative studies. In addition to demonstrating the existence of emotion/performance covariance, our studies showed that task difficulty (i.e., set-size) has an effect on the emotions participants experience during task execution. As expected, higher set-size resulted in increased unpleasant and decreased pleasant emotions during the task.

### VWM Tasks Induce Achievement Emotions

While there have been previous speculations about (predominantly negative) emotions occurring during typical lab-based cognitive performance tasks (Luck, 2014; Rouder et al., 2008), the present research was the first to systematically explore these assumptions. Our qualitative results from Study 1 revealed that participants experienced various different discrete emotions, both negative and positive in valence, during a VWM task, while overall negative emotions were mentioned more frequently than positive emotions. Interestingly, these emotional experiences not only varied between individuals, in that some participants experienced more joy and others more anger, but also within them: Participants experienced a range of varying emotions, while performing this laboratory task, for instance switching from joy, to boredom, to anger, and back to joy again during the course of one single experimental block. Importantly, though, joy, anger and frustration clearly were the most dominant emotional experiences reported by the participants in Study 1, whereas, by comparison, boredom was reported less frequently. This is an important finding of the present study in and of itself: While researchers in this field seemed to have been concerned with participant boredom, our results indicate that this seems to be less problematic than previously assumed, at least for the continuous color-report task examined here.

Beyond identifying which discrete emotions the participants experienced during the VWM task, we also employed qualitative content analysis to classify participants' think aloud utterances pertaining to potential reasons for their current emotional experiences. Across the various categories for triggers of emotions identified by this approach, one striking overarching theme emerged from this analysis: Participants seemed to be constantly aware that they could either succeed or fail at the VWM task (i.e., recalling the correct/exact color of the probed square versus failing to do so), and a clear majority of their thoughts centered around corresponding achievement appraisals. By implication, the predominant type of emotions participants experience during the examined VWM task seem to be achievement emotions (Pekrun, Lichtenfeld, et al., 2017).

It is worth noting that within the think-aloud transcripts, it became apparent that some participants hinted towards using certain strategies in order to enhance their performance on the VWM task such as verbalization of the colors (Souza & Skóra, 2017), grouping (Morey, 2019), and ensemble representations (e.g., Brady and Al-

varez, 2015; Liesefeld and Müller, 2019), and they expressed varying degrees of mind wandering (e.g., Robison and Unsworth, 2018). As this was not the focus of our research question, we did not further follow up on those observations. However, it may be interesting to explore in more detail how those phenomena are linked with task-induced emotions in future research. Furthermore, it is worth noting that some participants expressed frustration about the fact that they did not get performance feedback; thus they apparently felt that the task would be more satisfying if performance feedback was provided. Instead, we would actually expect that provision of performance feedback would intensify any emotional experiences during the task (positive after success feedback, negative after failure feedback), thus yet exacerbating the potential emotion—performance links. We therefore deliberately decided against providing any performance feedback in the present study to see whether even in that situation, participants would experience achievement emotions. Future research may explore effects of performance feedback on task-induced emotions during cognitive tasks, and potential corresponding effects on task performance and emotion—performance covariation.

When realizing that the continuous color-report task has such strong task-inherent achievement requirements, which trigger achievement emotions in the large majority of participants, it seems helpful to consider Pekrun's (2006) control-value theory of achievement emotions to better understand the possible underlying processes of emotion elicitation during VWM task performance. In this theory, Pekrun proposes that individuals vary in their emotional experiences depending on their appraisals concerning the achievement activity and its outcome, in terms of subjective control appraisals (which correspond with judgments of whether one can succeed at the task) and subjective value (which correspond with judgments of how important it is to succeed on the task). More specifically, Pekrun (2006, 2018) proposes that control appraisals determine the valence of emotions (e.g., enjoyment or pride in case of high control, anger, frustration or anxiety in case of low control), and value appraisals boost the emotional intensity (stronger with higher value). As such, the present findings imply that as long as participants truly commit to the task – i.e., they accept that selecting the correct color is important, and continually monitor their own performance by judging whether or not they think they got it right – both negative and positive achievement emotions are bound to occur during the task. Yet, this also implies that participants will vary in the levels of task-induced achievement emotions, depending on how much importance they attach to selecting the correct color, and how successful they sense they are at doing so.

## Large Array Sizes Increase Negative Task Emotions and Decrease Positive Task Emotions

A key finding of the present research was that emotional processes involved in performing a VWM task were influenced by task difficulty, with larger set size being associated with decreased positive and increased negative emotions. This was shown in Study 3 for the emotions joy, pride, anger and frustration. This finding was replicated using a larger sample in Study 4, where the effect of set-size on boredom also reached significance (more boredom for set size 8 than set size 4). The variation of set-sizes is a common practice by many researchers who use the continuous color-report task,

and our findings suggest that the emotions induced by varying set-sizes might be a confound that has not yet been sufficiently considered. Rouder et al. (2008) speculated that some participants may be intimidated by the harder trials, which in turn may harm their VWM performance. The present study is the first to provide evidence that participants experience more positive emotions and less negative emotions for smaller compared to larger set-sizes.

Again, control-value theory (Pekrun, 2006, 2018) provides the theoretical underpinning for explaining this finding: with larger set-sizes, the participants' control over succeeding at the task decreases. Given that a set-size of eight clearly exceeds the VWM capacity of a large proportion of participants, their chances of remembering the correct color are low, and thus their subjective appraisals of whether they can succeed at the task are bound to be poor. As a result, negative achievement emotions emerge. In contrast, easier trials (e.g., set-size 4) are appraised by the participants as more controllable, resulting in more pleasant task emotions. A key implication of the present study is thus that the task emotions affected by set-size represent an essential potential confound in any study that seeks to explore any effects of set-size in the context of VWM research.

## Task Emotions and VWM Performance are Systematically Linked

The second and overarching goal of the present research was to quantify the links between task-induced emotions and performance. Overall, our findings implied that there are positive links between pleasant emotions and VWM performance, and negative links between unpleasant emotions and VWM performance. These systematic emotion—performance links were demonstrated both across individuals (Study 2), and within individuals (Studies 3 and 4). It is worth noting that our attempt to conceptually replicate the Study 2 findings by adding an overall retrospective emotion rating at the very end of the experiment (in addition to the repeated emotion ratings within the experimental block) failed. We suppose that adding this question to the design of Study 4 did not provide the same measurement as obtained in Study 2, because performing similar ratings with regard to the current emotional experience multiple times throughout the experiment, unfortunately but not unexpectedly, affects how participants perform the final rating.

The emotion—performance links as demonstrated in our studies are consistent with previous research demonstrating that positive emotions are associated with enhanced working memory performance (e.g., Brose et al., 2014) and negative affect with decreased performance (Figueira et al., 2017). Importantly, these results are correlational and as such do not allow to draw any causal implications. Based on claims and corresponding findings from field studies in applied academic contexts (e.g., Pekrun, Lichtenfeld, et al., 2017), we find it plausible that emotions and performance are linked through reciprocal causation, meaning that both causal directions exist. Taking task enjoyment as an example, on the one hand, doing well on the task causes participants to experience joy, but on the other hand, enjoying the task also leads to participants doing well, as it boosts their task motivation and willingness to invest effort, and focuses their attention on the task.

As mentioned above, participants were not provided with any feedback on their task achievement, so they did not know for certain how well or how poorly they were doing. However, participants

seemed to have a good sense of their task achievement, as they often commented on their overall success rate and subjective success at individual trials in Study 1, which might indicate that they know when they have forgotten a probed item and therefore have to guess. Further, in Study 2, results showed the better participants performed relative to others, the more they reported enjoying the task and feeling proud, and the poorer they performed relative to others, the more they reported anger, frustration, and boredom. Studies 3 and 4 further confirmed that those emotion/performance links also emerged on the intra-individual level. In other words, the dynamics of participants' emotions across the course of the experiment, as assessed through multiple emotion ratings after short experimental sub-blocks, co-fluctuated with the dynamics of the performance across those sub-blocks, within the participants: the better the participants did at a certain point within the experimental block, the better they felt at this moment, while when they performed more poorly, negative emotions were aroused within them. In turn, assuming reciprocal causation, this correlative pattern also implies that the better participants felt during task execution, the higher they performed.

We propose that such reciprocal causation between task performance and task emotions should result, in case of task success, in upward spirals, and in the case of task failure, in downward spirals. Specifically, we propose that those participants who truly have higher capacity will quickly get a subjective feeling of doing well on the task, which makes them joyous. In turn, we suppose that this task enjoyment boosts their sense of challenge and opportunity to perform during task, and as a result, they do even better at it. In contrast, those participants who have lower capacity will quickly get a subjective feeling of doing poorly on the task, which makes them angry and frustrated. This anger and frustration will undermine their task performance. At best, they will keep trying, complying with the task requirement asking them to recall the correct color. However, they may also, for the sake of emotion regulation (c.f. Gross, 2002), re-appraise the situation and decide that doing well on the task is not so important for them. This then may result in decreased task commitment, which further undermines their performance.

We believe that there is an important implication of these possible reciprocally spiraling emotion-performance links which seem to be initiated during classical VWM tasks, such as the continuous color-report task (or even during cognitive tasks in general), due to their strong task-inherent achievement requirement. We suggest that thus-obtained capacity scores are dually biased due to the emotional processes just described: They are positively biased, the higher the true capacity, and negatively biased, the lower the true capacity, thus resulting in an overestimation of inter-individual variability in task performance (Vogel & Awh, 2008). This may not be the case for every participant, as individuals may vary in responding to the task-inherent achievement requirements, that is, in how much they value solving the task correctly. Future research may follow up on this notion we see implied by our findings by systematically exploring the emotional responses of low versus high achievers in VWM and how this influences VWM performance (see also Fukuda et al., 2010; Luck and Vogel, 2013).

In sum, the present results provide substantial evidence to confirm earlier speculations about VWM tasks inducing certain emotional experiences in participants. These task-induced emotions are systematically linked with VWM performance and this may be worth considering in future cognitive and VWM research. As researchers, we would like participants to be more like machines sometimes, so we can examine their "hardware" most accurately. However, it seems that human functioning is more complex and highly interacts with emotional experiences, so that future research needs to account for task-induced emotions.

## Context of Research

In an ongoing collaboration, we combine theories and findings of two very different fields of study within the same discipline: educational and general psychology. In particular, the reported study combines ACF's expertise in achievement emotions (e.g., Frenzel et al., 2018; Frenzel et al., forthcoming), a construct traditionally researched in applied academic settings, with HRL's expertise in visual working memory (e.g., Constant and Liesefeld, 2021; Liesefeld et al., 2020; Liesefeld and Müller, 2019), typically researched in basic lab contexts. By our interdisciplinary approach, we were able to gain novel insights into both constructs, which bear important implications for both basic visual working memory research and applied achievement emotion research. In future research, we aim to further inspect these implications for both fields of study.

## References

Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, *106*(1), 20–29. https://doi.org/10/cjfxwk

Beal, D. J., Weiss, H. M., Barros, E., & MacDermid, S. M. (2005). An episodic process model of affective influences on performance. *Journal of Applied Psychology*, *90*(6), 1054–1068. https://doi.org/10/b8q45n

Brady, T. F., & Alvarez, G. A. (2015). No evidence for a fixed object limit in working memory: Spatial ensemble representations inflate estimates of working memory capacity for complex objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(3), 921–929. https://doi.org/10/ggt428

Brose, A., Lövdén, M., & Schmiedek, F. (2014). Daily fluctuations in positive affect positively co-vary with working memory performance. *Emotion*, *14*(1), 1–6. https://doi.org/10/f22drg

Brun, G., Doğuoğlu, U., & Kuenzle, D. (Eds.). (2008). *Epistemology and emotions*. Ashgate.

Constant, M., & Liesefeld, H. R. (2021). Massive effects of saliency on information processing in visual working memory. *Psychological Science*, *32*(5), 682–691. https://doi.org/10/gjk9jh

Conway, A. R., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, *7*(12), 547–552. https://doi.org/10/c8jjx5

Csikszentmihalyi, M. (2000). *Beyond boredom and anxiety: Experiencing flow in work and play*. Jossey-Bass Publishers.

Deci, E. L., & Ryan, R. M. (1991). A motivational approach to self: Integration in personality. *Nebraska Symposium on Motivation, 1990: Perspectives on motivation* (pp. 237–288). University of Nebraska Press.

D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction*, *22*(2), 145–157. https://doi.org/10/fv2b3c

Dukes, D., Abrams, K., Adolphs, R., Ahmed, M. E., Beatty, A., Berridge, K. C., Broomhall, S., Brosch, T., Campos, J. J., Clay, Z., Clément, F., Cunningham, W. A., Damasio, A., Damasio, H., D'Arms, J., Davidson, J. W., de Gelder, B., Deonna, J., de Sousa, R., . . . Sander, D. (2021). The rise of affectivism. *Nature Human Behaviour*, *5*(7), 816–820. https://doi.org/10/gkhxms

Dweck, C. S., & Elliott, E. S. (1983). Achievement motivation. In M. Hetherington (Ed.), *Socialization, personality, and social development* (pp. 643–691). Wiley.

Elliot, A. J., Murayama, K., & Pekrun, R. (2011). A 3 × 2 achievement goal model. *Journal of Educational Psychology*, *103*(3), 632–648. https://doi.org/10/ch6d89

Figueira, J. S. B., Oliveira, L., Pereira, M. G., Pacheco, L. B., Lobo, I., Motta-Ribeiro, G. C., & David, I. A. (2017). An unpleasant emotional state reduces working memory capacity: Electrophysiological evidence. *Social Cognitive and Affective Neuroscience*, *12*(6), 984–992. https://doi.org/10/f9zrwj

Frenzel, A. C., Becker-Kurz, B., Pekrun, R., Goetz, T., & Lüdtke, O. (2018). Emotion transmission in the classroom revisited: A reciprocal effects model of teacher and student enjoyment. *Journal of Educational Psychology*, *110*(5), 628–639. https://doi.org/10/ggb623

Frenzel, A. C., Goetz, T., & Loderer, K. (forthcoming). Emotions and emotion regulation. In P. Schutz & K. R. Muis (Eds.), *Handbook of Educational Psychology* (4th ed.). Routledge/Taylor-Francis.

Fukuda, K., Vogel, E., Mayr, U., & Awh, E. (2010). Quantity, not quality: The relationship between fluid intelligence and working memory capacity. *Psychonomic Bulletin & Review*, *17*(5), 673–679. https://doi.org/10/dwkkc6

Gergelyfi, M., Jacob, B., Olivier, E., & Zénon, A. (2015). Dissociation between mental fatigue and motivational state during prolonged mental activity. *Frontiers in Behavioral Neuroscience*, *9*. https://doi.org/10/ghs9ww

Goetz, T., & Hall, N. C. (2013). Emotion and achievement in the classroom. *International guide to student achievement* (pp. 192–195). Routledge/Taylor & Francis Group.

Gross, J. J. (2002). Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology*, *39*(3), 281–291. https://doi.org/10/dhr5xt

Hopstaken, J. F., van der Linden, D., Bakker, A. B., & Kompier, M. A. J. (2015). A multifaceted investigation of the link between mental fatigue and task disengagement: Mental fatigue and task disengagement. *Psychophysiology*, *52*(3), 305–315. https://doi.org/10/f63dqp

Hopstaken, J. F., van der Linden, D., Bakker, A. B., Kompier, M. A. J., & Leung, Y. K. (2016). Shifts in attention during mental fatigue: Evidence from subjective, behavioral, physiological, and eye-tracking data. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(6), 878–889. https://doi.org/10/f8qpsg

Johnson, M. K., McMahon, R. P., Robinson, B. M., Harvey, A. N., Hahn, B., Leonard, C. J., Luck, S. J., & Gold, J. M. (2013). The relationship between working memory capacity and broad measures of cognitive ability in healthy adults and people with schizophrenia. *Neuropsychology*, *27*(2), 220–229. https://doi.org/10/f4sct8

Kanai, R., & Rees, G. (2011). The structural basis of inter-individual differences in human behaviour and cognition. *Nature Reviews Neuroscience*, *12*(4), 231–242. https://doi.org/10/cv46sf

Kyllingsbæk, S., & Bundesen, C. (2009). Changing change detection: Improving the reliability of measures of visual short-term memory capacity. *Psychonomic Bulletin & Review*, *16*(6), 1000–1010. https://doi.org/10/c32bch

LaBar, K. S., & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews Neuroscience*, *7*(1), 54–64. https://doi.org/10/bk9dcz

Lagattuta, K. H., & Thompson, R. A. (2007). The development of self-conscious emotions: Cognitive processes and social influences. *The self-conscious emotions: Theory and research* (pp. 91–113). Guilford Press.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). *International Affective Picture System (IAPS): Technical manual and affective ratings*. https://doi.org/10/hcrv

Laybourn, S., Frenzel, A. C., Constant, M., & Liesefeld, H. R. (2020). Unintended Emotions in the Laboratory: Emotions Incidentally Induced by a Standard Visual Working Memory Task Relate to Task Performance. https://osf.io/dr62j/

Liesefeld, H. R., Liesefeld, A. M., Sauseng, P., Jacob, S. N., & Müller, H. J. (2020). How visual working memory handles distraction: Cognitive mechanisms and electrophysiological correlates. *Visual Cognition*, *28*(5-8), 372–387. https://doi.org/10/gg5vsv

Liesefeld, H. R., & Müller, H. J. (2019). Current directions in visual working memory research: An introduction and emerging insights. *British Journal of Psychology*, *110*(2), 193–206. https://doi.org/10/gfvm2p

Long, F., Ye, C., Li, Z., Tian, Y., & Liu, Q. (2020). Negative emotional state modulates visual working memory in the late consolidation phase. *Cognition and Emotion*, *34*(8), 1646–1663. https://doi.org/10/ghp84s

Luck, S. J. (2014, May 30). *An introduction to the Event-Related Potential technique* (2nd ed.). A Bradford Book.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281. https://doi.org/10/cqwttd

Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: From psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, *17*(8), 391–400. https://doi.org/10/39v

Luria, R., Balaban, H., Awh, E., & Vogel, E. K. (2016). The contralateral delay activity as a neural measure of visual working memory. *Neuroscience & Biobehavioral Reviews*, *62*, 100–108. https://doi.org/10/f8dp5m

Mayring, P. (2014). *Qualitative content analysis: Theoretical foundation, basic procedures and software solution*.

Mayring, P., & Fenzl, T. (2014). *QCAmap. A software for qualitative content analysis*. https://www.qcamap.org/

Morey, C. C. (2019). Perceptual grouping boosts visual working memory capacity and reduces effort during retention. *British Journal of Psychology*, *110*(2), 306–327. https://doi.org/10/ghs9vm

Muis, K. R., Pekrun, R., Sinatra, G. M., Azevedo, R., Trevors, G., Meier, E., & Heddy, B. C. (2015). The curious case of climate change: Testing a theoretical model of epistemic beliefs, epistemic emotions, and complex learning. *Learning and Instruction*, *39*, 168–183. https://doi.org/10/ggd6zw

Muis, K. R., Psaradellis, C., Lajoie, S. P., Di Leo, I., & Chevrier, M. (2015). The role of epistemic emotions in mathematics problem solving. *Contemporary Educational Psychology*, *42*, 172–185. https://doi.org/10/gnv4rg

Mulligan, K., & Scherer, K. R. (2012). Toward a Working Definition of Emotion. *Emotion Review*, *4*(4), 345–357. https://doi.org/10/gf7nr3

Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, *44*(4), 369–378. https://doi.org/10/d3h3b6

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/d6rrq9

Pekrun, R. (2018). Control-value theory: A social-cognitive approach to achievement emotions. In G. A. D. Liem & D. M. McInerney (Eds.), *Big theories revisited 2: A volume of research on sociocultural influences on motivation and learning.* Information Age Publishing.

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology*, *36*(1), 36–48. https://doi.org/10/djzhcs

Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K., & Goetz, T. (2017). Achievement Emotions and Academic Performance: Longitudinal Models of Reciprocal Effects. *Child Development*, *88*(5), 1653–1670. https://doi.org/10/gbw6vg

Pekrun, R., & Perry, R. P. (2014). Control-value theory of achievement emotions. *International handbook of emotions in education* (pp. 120–141). Routledge/Taylor & Francis Group.

Pekrun, R., & Stephens, E. J. (2012). Academic emotions. In K. R. Harris, S. Graham, T. Urdan, S. Graham, J. M. Royer, & M. Zeidner (Eds.), *APA educational psychology handbook, Vol 2: Individual differences and cultural and contextual factors.* (pp. 3–31). American Psychological Association. https://doi.org/10/bdqx6d

Pekrun, R., Vogl, E., Muis, K. R., & Sinatra, G. M. (2017). Measuring emotions during epistemic activities: The Epistemically-Related Emotion Scales. *Cognition and Emotion*, *31*(6), 1268–1276. https://doi.org/10/ghphct

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, *17*(03). https://doi.org/10/d6f7tr

Putwain, D. W., Becker, S., Symes, W., & Pekrun, R. (2018). Reciprocal relations between students' academic enjoyment, boredom, and achievement over time. *Learning and Instruction*, *54*, 73–81. https://doi.org/10/gdbvxc

Robison, M. K., & Unsworth, N. (2018). Cognitive and contextual correlates of spontaneous and deliberate mind-wandering. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(1), 85–98. https://doi.org/10/gcx8th

Rouder, J. N., Morey, R. D., Cowan, N., Zwilling, C. E., Morey, C. C., & Pratte, M. S. (2008). An assessment of fixed-capacity models of visual working memory. *Proceedings of the National Academy of Sciences*, *105*(16), 5975–5979. https://doi.org/10/fngr7w

Souza, A. S., & Skóra, Z. (2017). The interplay of language and visual perception in working memory. *Cognition*, *166*, 277–297. https://doi.org/10/gbm62h

Souza, A. S., Thaler, T., Liesefeld, H. R., Santos, F. H., Peixoto, D. S., & Albuquerque, P. B. (2021). No evidence that self-rated negative emotion boosts visual working memory precision. *Journal of Experimental Psychology: Human Perception and Performance*, *47*(2), 282–307. https://doi.org/10/ghwzjb

Spachtholz, P., Kuhbandner, C., & Pekrun, R. (2014). Negative affect improves the quality of memories: Trading capacity for precision in sensory and working memory. *Journal of Experimental Psychology: General*, *143*(4), 1450–1456. https://doi.org/10/f6c5dp

van Someren, M. W., Barnard, Y. F., & Sandberg, J. A. C. (1995). The think aloud method: A practical guide to modelling cognitive processes. *Information Processing & Management*, *31*(6), 906–907. https://doi.org/10/fdzb2d

Vogel, E. K., & Awh, E. (2008). How to exploit diversity for scientific gain: Using individual differences to constrain cognitive theory. *Current Directions in Psychological Science*, *17*(2), 171–176. https://doi.org/10/bcd74p

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*(6984), 748–751. https://doi.org/10/dpb3j5

Vogl, E., Pekrun, R., Murayama, K., & Loderer, K. (2020). Surprised–curious–confused: Epistemic emotions and knowledge exploration. *Emotion*, *20*(4), 625–641. https://doi.org/10/gjp858

Vogl, E., Pekrun, R., Murayama, K., Loderer, K., & Schubert, S. (2019). Surprise, curiosity, and confusion promote knowledge exploration: Evidence for robust effects of epistemic emotions. *Frontiers in Psychology*, *10*, 2474. https://doi.org/10/gngt82

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070. https://doi.org/10/ck3

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision*, *4*(12), 11. https://doi.org/10/bkrxvz

Xie, W., & Zhang, W. (2016). Negative emotion boosts quality of visual working memory representation. *Emotion*, *16*(5), 760–774. https://doi.org/10/f848bj

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233–235. https://doi.org/10/c693sk

## Supplemental Material

## Information on Transcription of "Think Aloud" Protocols

1. Use "P:" or "R:", respectively, to indicate that the participant/researcher is talking

2. Transcribe verbatim, except when the participant recites colors. Summarize this by adding the following to the transcription: "recites colors"

3. Pauses in participants' monologues are not noted

4. Fillers such as "erm" do not need to be transcribed

5. Include non-verbal utterances, such as laughing or coughing, in brackets and in italics

6. Replace names of people, towns or any other such information with "XYZ"

## Coding Scheme

Data regarding researcher questions (RQ) 1a and 1b were analyzed using and inductive procedure (Mayring, 2014), as, to the authors' knowledge, there are no prior findings regarding these research questions. Therefore, categories needed to be extracted from the textual material itself on the basis of a predefined selection criterion and level of abstraction (see Table S1 and Table S2).

## RQ1a: What do participants feel when performing a VWM task?

*Selection Criterion.* Select all text passages in which participants referred to any sort of emotional states experienced during the VWM task. Do not select any references to emotional states that pertain to the think-aloud task (e.g., "It stresses me out to talk and perform the task at the same time").

*Level of Abstraction.* Categories are formulated as specific emotional states or personal feelings, which participants referred to during the VWM task.

## RQ1b: Which aspects of the VWM task trigger these discrete emotions?

*Selection Criterion.* Select all text passages in which participants explicitly referred to or hinted at reasons, sources or processes related to the emotional states or feelings experienced/perceived during the VWM task.

*Level of Abstraction.* Categories are formulated as specific reasons, sources or processes affecting or leading to participants' emotional states and feelings experienced/perceived during the VWM task.

**Table S1**

*Coding Scheme for RQ1a*

| Category Definition | Anchor Examples | Coding Guidelines |
|---|---|---|
| Anger | That just makes me angry. | Clear meaning component in the text; Multiple responses allowed (applies to all categories here) |
| Frustration | It's really frustrating | |
| Joy | I'm enjoying this; I was happy about that | |
| Boredom | It's always the same thing; It's starting to get boring | |
| Tension/Nervousness | It's probably because of nerves. | |
| Confusion | I'm a bit confused. | |
| Desperation | I get desperate | |
| Hope | I think it's better now, I hope I'll do better. | |
| Shame | I'm a bit ashamed. | |
| Disappointment | I'm disappointed when I don't know it (the correct color). | |
| Uncertainty | I feel uncertain and that's not pleasant. | |
| Anxiety | I feel afraid of not doing it right. | |

**Table S2**

*Coding Scheme for RQ1b*

| Category Definition | Anchor Examples | Coding Guidelines |
|---|---|---|
| Self-expectations | This isn't really that difficult. | Clear meaning component in the text; Multiple responses allowed (applies to all categories here) |
| VWM task is challenging (negative sense) | I'm out of my depth. | |
| General judgement of the VWM task | At the end of the day, it's just like a game. | |
| Dissatisfaction with the VWM task design | It's always the same thing. | |
| Change in motivation | It doesn't matter to me that much anymore. | |
| Social comparison | I'm always asking myself, if I'm that bad or if the others are also this bad. | |
| Referring to missing performance feedback | It would be interesting to know your score. | |
| Referring to missing time reference | You don't know when it'll be over. | |
| VWM task is challenging (positive sense) | I want to continue doing this and I want to do well. | |
| Thoughts on strategies to improve achievement | You start and build themes. | |