Domain-specific working memory loads selectively increase negativity bias

|  |
| --- |
| Nicholas R. Harp1 & Maital Neta1 |
| 1 University of Nebraska-Lincoln |
|  |

# Author note

Correspondence concerning this article should be addressed to Nicholas R. Harp. E-mail: [nharp@huskers.unl.edu](mailto:nharp@huskers.unl.edu)

**Abstract**

Cognitive resources are required for successful executive functioning. When resources needed for a particular task are limited, such as under cognitive load, then task performance is impaired. When faced with emotional ambiguity (i.e., stimuli that do not convey a clear positive or negative meaning), our decisions to approach or avoid appear to rely heavily on domain-specific emotion resources. Specificially, previous work has shown that the default response to emotional ambiguity is negative, and that positive categorizations are thought to require top-down regulatory resources. Here, we investigated the effects of domain-specific (non-emotional vs. emotional) and domain-general (low versus high) load on evaluations of emotional ambiguity (i.e., surprised facial expressions). As predicted, a mixed effects model revealed domain-specific effects at both low (p = .001) and high (*p* < .001) load, such that ratings of surprise were more negative during any emotional load. Consistent with prior work, analyses of response trajectories revealed that, under low load (regardless of domain), there was greater response competition on positive than negative trials, showing that positive categorizations are characterized by an initial attraction to negativity. However, under high load, response competition increased for negative trials (p = .005). These results suggest domain-specific load depletes the regulatory resources necessary for positive categorizations of emotional ambiguity, but that greater domain-general load impacts response competition that weighs in across the categorization process.

Keywords: [up to 5 words / phrases]

Domain-specific working memory loads selectively increase negativity bias

The availability of cognitive resources is necessary for successfully navigating our daily lives, including for adaptive processes in attention deployment (Franconeri, Alvarez, & Cavanagh, 2013), planning (Hayes-Roth & Hayes-Roth, 1979; Kliegel, Martin, McDaniel, & Phillips, 2007), decision-making (Deck & Jahedi, 2015; Whitney, Rinehart, & Hinson, 2008), inhibition (Ward & Mann, 2000) and cognitive control (Deveney & Pizzagalli, 2008). Thus, engagement in one task that taps a cognitive resource (i.e., cognitive load) impairs performance on a concomitant or subsequent task tapping the same resource (Richeson & Trawalter, 2005; Baumeister & Heatherton, 1996; Franconeri, Alvarez, & Cavanagh, 2013; Kahneman, 1973; Storbeck, 2012). For example, imagine a student attending a lecture while also text messaging a friend. As the student considers how to respond in their next message and directs cognitive resources (attention) towards the conversation and away from the lecture, the student’s ability to understand and remember the lecture material will suffer. Directing cognitive resources between different tasks in this manner taxes an already limited pool of cognitive resources (Baumeister & Heatherton, 1996; Kahneman, 1973). And on a larger scale, the accumulation of cognitive depletion can have widespread societal implications (e.g., burnout and absenteeism; Diestel & Schmidt, 2011). For instance, in emotionally demanding occupations (e.g., healthcare positions), cognitive depletion is associated with worse job performance (Ihle, Borella, Rahnfeld, Müller, Enge, Hacker, Wegge, Oris, & Kliegel, 2015; Motowidlo, Packard, & Manning, 1986) and increased job-related stress that has adverse downstream effects on executive functioning (Privitera, Rosenstein, Plessow, & LoCastro, 2014; Starcke, Wiesen, Trotzke, & Brand, 2016).

Notably, the domain of the resource depletion matters; that is, if the resources required for a particular task are not depleted, then task performance may be unaffected by the load. In other words, if it were the case that resources taxed by a load were not required to complete a task, then there may be no behavioral consequences on one’s task performance. For instance, when asked to remember the emotional expression of a face, rather than its identity, participants were less accurate on subsequent decisions in pairing concepts according to emotional properties (e.g., couple-happy) compared to perceptual properties (e.g., lemon-yellow; Vermeulen, Niedenthal, Pleyers, Bayot, & Corneille, 2014). Neuroimaging research extends these findings by suggesting separable effects of load as a function of the domain (non-emotional versus emotional). For instance, changing the nature of cognitively demanding tasks, such that participants attend to and remember an emotional component (e.g., emotional expression instead of an identity) of the same stimuli, results in the recruitment of dissociable neural resources (Egner, Etkin, Gale, & Hirsch, 2008; Neta & Whalen, 2011). Therefore, when resources are engaged with an emotional load, the resources are no longer available for other emotional processes, and performance on these tasks will likely be impaired. As such, the effect of load on a concurrent or subsequent task may illuminate the processes required for that task.

**Resources used for categorizing emotional ambiguity**

While humans readily make judgments about others with only limited information and resources (e.g., judging trustworthiness, attractiveness, and emotion; Bar, Neta, & Linz, 2006; Said & Todorov, 2011; Todorov, Baron, & Oosterhof, 2008; Cloutier, Heatherton, Whalen, & Kelley, 2008; Brooks, Chikazoe, Sadato, & Freeman, 2019; Carroll & Russell, 1996), some judgments are vulnerable to resource depletion. For example, we swiftly categorize new information, including facial expressions, by valence, which is crucial for guiding social behavior (e.g., approach-avoidance; Krieglmeyer, Deutsch, De Houwer, & De Raedt, 2010; and group membership or affiliation; Taskhay & Rule, 2015; Taskhay & Rule, 2018). Although some facial expressions are easily categorized as positive (happy) or negative (angry), others (surprise) require more resources due to the nature of their valence ambiguity (Neta et al., 2009; Neta & Tong, 2016; Petro, Tong, Henley, & Neta, 2018). Indeed, surprised expressions can predict both positive (e.g., winning the lottery) and negative (e.g., a car accident) outcomes. Thus, there are individual differences in the tendency to categorize surprised faces as having a more positive or negative meaning, which is known as one’s *valence bias* (Neta, Kelley, & Whalen, 2013; Neta et al., 2009; Neta & Whalen, 2010).

Despite the individual differences in valence bias, there appears to be an initial negativity across people (i.e., *initial negativity hypothesis*; Neta, Davis, & Whalen, 2011; Neta et al., 2009; Neta & Whalen, 2010; Petro et al., 2018). Under this framework, positive categorizations rely on additional emotion regulation resources that help to override the initial negativity. Support for this model comes from studies using mousetracker (Freeman & Ambady, 2010), which offers a rich insight into the process underlying decision-making and associated response competition (Calcagni, Lombardi, & Sulpizio, 2017; Freeman, Dale, & Farmer, 2011; Hehman, Stolier, & Freeman, 2015). This work revealed that trajectories to the negative response option are more direct, whereas positive categorizations are characterized by greater response competition for the unselected (negative) response (Brown et al., 2017; Mattek et al., 2016). Neuroimaging work has also supported this initial negativity hypothesis by demonstrating that the amygdala, which responds to more bottom-up signals of emotion (Aggleton and Saunders, 2000; Adolphs, 1994; LeDoux, 2000), and the ventromedial prefrontal cortex (vmPFC), a putative top-down regulatory region (Motzkin, Philippi, Wolf, Baskaya, & Koenigs, 2015), show inverse activity that is predictive of subjective categorizations of surprised expressions (Kim et al., 2003). Specifically, more frequent negative categorizations are associated with greater amgydala activity (Kim et al., 2003; Neta & Whalen, 2010) and positive categorizations are associated with greater vmPFC activity (Kim, Somerville, Johnstone, Alexander, & Whalen, 2003). More recently, Petro, Tong, Henley, & Neta (2018) found that participants with a more positive valence bias showed greater surprise-related activity in brain regions recruited during an explicit emotion regulation (cognitive reappraisal) task. Taken together, positive categorizations appear to rely on more regulatory resources than negative categorizations, therefore concurrent demands that use those same resources are likely to interfere with one’s ability to see ambiguity in a positive light.

Although some previous work found no effect of load on categorizations of surprised faces (Mattek et al., 2016), this study used only non-emotional load (i.e., remembering a number sequence). Thus, it could be that a domain-specific (emotional) load will more effectively deplete the resources putatively required for a positive bias, resulting in more negative categorizations. It is worth noting that although the non-emotional load did not affect the categorizations, it did impact the response trajectories such that high load diminished attraction towards modal responses (i.e., in line with one’s bias) during categorizations of surprised faces (Mattek et al., 2016). Thus, we predict that domain-general load (i.e., load that is irrespective of domain) will be associated with a similar change in the response trajectories, particularly diminishing the difference in response competition between negative and positive categorizations.

## The present study

In the present study we tested the effect of cognitive load on valence bias, as a function of load (low versus high) and domain (non-emotional versus emotional). To do this, we manipulated the amount of material that participants needed to remember and the domain of that material while concurrently categorizing the valence of facial expressions. First, we predict that there will be no effect of load on categorizations of surprised faces, replicating Mattek and colleagues (2016). However, we do expect to find an effect of domain on categorizations, such that an emotional load will result in more negative categorizations than a non-emotional load, suggesting that emotional load depletes the resources required for seeing ambiguity in a positive light. Further, we predict greater response competition for positive than negative categorizations under low load, consistent with previous work (Brown et al., 2017; Mattek et al., 2016), and that this effect will be mitigated under high load (Mattek et al., 2016), irrespective of the load domain.

# Methods

## Participants

Fifty-nine participants (*M*age = 19.03 years, SD = 1.70 years, 49 female) were recruited from the undergraduate research pool at the University of Nebraska-Lincoln. The data from nine participants were excluded due to technical difficulties that prevented data from being saved. The final sample included the remaining 50 participants (*M*age = 18.82 years, SD = 1.19 years, 41 female), and all identified as White/Caucasian without Hispanic/Latinx ethnicity. All subjects provided written informed consent in accordance with the Declaration of Helsinki and all procedures were approved by the University of Nebraska-Lincoln Institutional Review Board (Approval #20141014670EP). Each participant received course credit for completing the study.

## Procedure

The task was conducted using MouseTracker software (Freeman & Ambady, 2010) and was structured to closely resemble the cognitive load task used by Mattek, Whalen, Berkowitz, and Freeman (2016), which used a single digit (low load) or seven digit sequence (high load). The trials were self-initiated; that is, the participant initiated each trial at their own pace by clicking the “start” button at the bottom of the screen. After initiating the trial, a fixation cross appeared (1000 ms), then participants viewed an image matrix consisting of 2 or 6 images (low or high load, respectively) with either emotional or non-emotional properties (equal number of trials) for 4000 ms (Figure 1). Participants were instructed to remember these images for the duration of the trial (i.e., until the memory probe at the end of the trial). After the image matrix, a happy, angry, or surprised face appeared for 1000 ms, and the participants categorized the face as positive or negative using the computer mouse. Finally, a single image probe appeared (5000 ms), and participants used the computer mouse to indicate whether the image probe was present in the previous image matrix by clicking either yes (i.e., the image was present) or no (i.e., the image was not present). The experimenter guided participants through a practice face rating and memory probe trial, after which they completed a total of 144 trials, and their mouse movements were recorded throughout. Notably, in two-choice designs, maximum deviations are often conceptualized as a measure of response competition for ultimately unchosen responses (Calcagni, Lombardi, & Sulpizio, 2017; Freeman, Dale, & Farmer, 2011; Hehman, Stolier, & Freeman, 2015).

## Stimuli

A total of 288 scenes (72 positive, 72 negative, and 144 neutral) were selected from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2008) for use in the matrices. An additional 61 IAPS images were used during memory probes, but never appeared in the image matrices. The positive and negative images used in the matrices did not differ in arousal after testing with a Wilcoxon signed-rank test (*Z* = -0.23, *p* = 0.82). For the matrices with emotional properties, there were an equal number of positive and negative images within a matrix in order to avoid priming effects on the subsequent face ratings (e.g., Flexas, Rosselló, Christensen, Nada, La Rosa, & Munar, 2013), particularly given that categorizations of surprised faces are sensitive to valence priming (Neta et al., 2011).

The face stimuli included images from the NimStim (Tottenham et al., 2009) and Karolinska Directed Emotional Faces (Lundqvist, Flykt, & Öhman, 1998) stimuli sets, as in previous work (Brown et al., 2017; Neta & Whalen, 2010). The faces consisted of 34 unique identities including 12 angry, 12 happy, and 24 surprised expressions organized pseudorandomly.

A screenshot of a cell phone

Description automatically generated

## Figure 1: Example of a single trial. At the start of each trial, a fixation cross appeared (1000 ms), followed by an image matrix consisting of 2 (low load) or 6 (high load) images with either non-emotional or emotional properties. Participants were instructed to remember these images for the duration of the trial. Then, a happy, angry, or surprised face appeared for 1000 ms, which the participants were instructed to categorize as positive or negative using the computer mouse. Finally, a single image probe appeared (5000 ms), and participants used the computer mouse to indicate whether the image probe was present (yes) or not (no) in the previous image matrix.

## Data analysis

We used R (Version 3.6.0; R Core Team, 2019) for all our analyses. Data preprocessing, analysis, and plotting were completed in R using the mousetrap (Kieslich, Henninger, Wulff, Haslbeck, Schulte-Mecklenbeck, 2019), lme4 (Bates, Maechler, Bolker, & Walker, 2015), TOSTER (Lakens, 2017), diptest (Maechler, 2016), and ggplot2 (Wickham, 2016) packages. While it is possible that trials in which participants responded incorrectly to the memory probe indicated a manipulation failure (i.e., the participant was not maintaining the images in memory), we included all trials regardless of accuracy due to the lack of an objective method for determining whether or not the participants were at least attempting to remember the images in the matrix. Our primary dependent measures were valence bias, which is calculated as percent negative ratings for surprised faces across all trials, and maximum deviation, or the extent to which a response trajectory deviated or was attracted to the competing – unselected – response option. For the main test of our hypotheses, we compared the valence bias for the different working memory load conditions (high and low load, emotional and non-emotional load), and explored the effects of condition and trialwise categorizations (positive and negative) on maximum deviation.

In order to account for the interdependence among measurements due to the repeated measures design, we used a mixed effects modeling approach. Unlike the repeated measures ANOVA, mixed effects models can account for missing data in repeated measures designs, which was a concern in our analyses given that some participants rated surprise as either negative or positive on all trials (i.e., there were missing values for the analysis of surprise using trial-by-trial rating as a factor). Mixed effects approaches are an extension of ordinary least squares (OLS) regressions that also include both fixed and random effects. The interpretations of fixed effects follow the conventions of OLS regression, where the slope describes the average effect for each one unit increase in the predictor, while random effects allow the model to fit effects which vary randomly across the nested structure of the data. In other words, random effects allow the partitioning of variance to either lower or higher level units of analysis. In this analysis, the lower level units are the repeated measurements and the higher level units are the participants. Allowing the intercept to vary randomly as a function of the higher level units accounts for individual differences in bias or other measures, thus allowing better estimation of the working memory condition effects at the lower level of analysis: the repeated measurements. Additional justification for the mixed effects modeling approach comes from tests of statistical dependency among the measures for any given subject, which was revealed through an intraclass correlation (ICC) of .75 for ratings of surprised faces and .17 for maximum deviations.

To test the effects of experimental conditions (load: high versus low, and domain: emotional versus non-emotional) on ratings and maximum deviations, we used a linear mixed model with a Gaussian error distribution. This approach demonstrated better model fit than alternative options (i.e., gamma distribution), and is robust to violations of normality (Knief & Forstmeier, 2018) evidenced in our data by Shapiro-Wilks tests (*p*’s < .001). All model building was completed using full information maximum likelihood estimation to account for any missing data (e.g., if a participant did not rate any surprised faces as positive or negative).

# Results

## Subjective categorizations of ambiguity

First, a random intercept-only model was tested and the likelihood ratio test results supported the decision to model the intercept randomly across individuals (*p* < .001). This suggests there was individual variance in valence bias at baseline (i.e., even under low, non-emotional loads) that is best modeled as a separate intercept for each subject.Next, fixed effects for Load (low versus high), Domain (emotional versus non-emotional), and their interaction were added to the model. There was a significant effect of Domain, such that categorizations of surprised faces following an emotional load (M = 74.0, S.E. = 3.49) were more negative than those following non-emotional loads (M = 65.0, S.E. = 3.49; ß = 9.03, *t*(153) = 5.17, *p* < .001). There was no significant effect of Load (low load: M = 68.9, S.E. = 3.49; high load: M = 70.2, S.E. = 3.49; ß= 1.34; *t*(153) = .77, *p* = .45), nor was there a significant Domain × Load interaction (ß30 = 1.93, , S.E. = 3.46, *t*(150) = .56, *p* = .58).

Given the non-significant effect of Load in the model, we next completed equivalence testing (Lakens, Scheel, & Isager, 2018) for the low and high load conditions to assess whether there was evidence to support an effect size statistically equivalent to zero. The smallest effect size of interest for the present study was determined to be d = .22 or a raw score difference of 5.9% negativity between the conditions, which a post-hoc power analysis suggests the current study was powered at 33% to detect. The result of the equivalence test (*t*(197.69) = 1.22, *p* = 0.112) was non-significant, and thus does not support the claim that the effect of Load is equivalent to zero.

**Mixed Model:** Percent Negative Ratingsti = (β00 + r0i) + β10\*(Domainti) + β20\*(Loadti) β30\*(Domainti \* Loadti) + eti

![A screenshot of a cell phone

Description automatically generated]()

**Figure 2: Percent negative ratings across conditions.** Ratings during emotional loads were more negative than ratings during non-emotional loads ( *t*(153) = 5.17, *p* < .001), but there was no effect of load on ratings ( *t*(153) = .77, *p* = .45). Error bars represent the standard error of the mean. \*\*p ≤ .001.

**Maximum deviation**

Next, we examined the effect of our experimental manipulation and surprise categorizations (positive versus negative trials) on maximum deviation. First, a random intercept-only model was tested for maximum deviation, and a likelihood ratio test supported this decision to model the intercept randomly (*p* < .001). This means that individuals differed in their average maximum deviations at baseline (i.e., under low, non-emotional loads), and that the best fitting model includes an intercept for each subject individually.Next, fixed effects of Load (low versus high), Domain (emotional versus non-emotional), Rating (positive versus negative categorizations of surprise), and their interactions were added to the model. A significant Load × Rating interaction (β50 = .32, *t*(314) = 3.55, S.E. = .09, *p* < .001; Figure 3) revealed that, as expected, positive categorizations (M = .50, S.E. = .04) had larger maximum deviations than negative (M = .30, S.E. = .04; *t*(325) = 4.39, S.E. = .05, *p* < .001; Bonferroni corrected significance for these analyses *p* < .013) on low load trials. However, this difference was not present on high load trials (positive: M = .41, S. E. = .04; negative: M = .42, S. E. = .04; *t*(327) = -.31, S.E. = .05, *p* = .76), supporting our hypothesis that high load would diminish the attraction towards the ‘negative’ response on response trajectories during categorizations under high load. Specifically, maximum deviations for negative trials were larger on high load compared to low load trials (*t*(320) = -2.81, S.E. = .05, *p* = .005), and there was a trend for smaller maximum deviations for positive trials on high load compared to low load trials (*t*(324) = 1.93, p = .055).

**Mixed Model:** Maximum Deviationti = (β00 + r0i) + β10\*(Loadti) + β20\*(Domainti) + β30\*(Ratingti) + β40\*(Loadti)\*(Domainti) + β50\*(Loadti)\*(Ratingti) + β60\*(Domainti)\*(Ratingti) + β70\*(Loadti)\*(Domainti)\*(Ratingti) + eti



**Figure 3: Response competition as a function of load and categorization.** There was greater response competition, operationalized as maximum deviation, for positive than negative categorization, but only under low load (*t*(325) = 4.39, S.E. = .05, *p* < .001). Error bars represent the standard error of the mean. +p < .06, \*p < .05, \*\*p ≤ .001.

The increase in maximum deviation for negative categorizations under high load could result from either an increased attraction towards the competing (positive) response option that weighs in across the decision-making process or a qualitative change in the trajectories (e.g., erratic “flip-flopping” trajectories) that could be adjudicated by examining the modality of the response trajectories (Freeman & Dale, 2013). Specifically, multimodality would suggest multiple subpopulations of trajectory patterns (flip-flopping) at play. Hartigan’s dip statistic (HDS; Hartigan & Hartigan, 1985) was calculated for each trial condition, and revealed no evidence of multimodality for the ‘positive’ and ‘negative’ categorizations in both the low (positive: HDS = .03, *p* = .98; negative: HDS = .02, *p* = .99) and high load conditions (positive: HDS = .04, *p* = .68; negative: HDS = .03, *p* = .80).

**Memory probe accuracy**

We examined accuracy on the memory probe to assess differences in task difficulty. While accuracy on the probes was high across all trials (94.41%), there were differences as a function of Load and Domain (Table 1). Given that the ICC was 0, the random intercept model had singular fit, and likelihood ratio tests did not suggest any benefit to modeling the intercept randomly (*p* > .999, which is likely due to a large proportion of the data having the same value of 100% correct), a mixed effects model could not be used to examine the effects on on accuracy. As such, we used a repeated measures ANOVA to examine differences in memory probe accuracy as a function of the experimental conditions, but note that caution is warranted in interpretations of the model given the undesirable structure of the data (i.e., lack of variability, strong non-normality).

There was a significant effect of Load (*F*(1, 49) = 50.28, *p* < .001) andDomain (*F*(1, 49) = 10.49, *p* = .002), such that accuracy was higher for low than high load, and for non-emotional than emotional load. Further, there was a significant Load × Domain interaction (*F*(1, 49) = 11.06, *p* = .002), such that Domain had no significant effect in low load conditions (*t*(96) = .44, *p* = .661; Bonferroni corrected significance for these analyses p < .013), but under high load, accuracy was higher for non-emotional than emotional load (*t*(96) = -4.63, *p* < .001). Further, the effect of Load did not survive Bonferroni correction on non-emotional trials (*t*(95) = -1.99, *p* = .05), but on emotional trials, accuracy was higher on low load than high load trials (*t*(95) = -7.10, *p* < .001).

**Table 1: Descriptive statistics for memory probe accuracy across all conditions.**

|  |  |  |
| --- | --- | --- |
|  | **Condition** | **Mean (SD)** |
| **Emotional load** | **Low load** | **98.83% (.04)** |
| **High load** | **88.33% (.11)** |
| **Non-emotional load** | **Low load** | **98.18% (.05)** |
| **High load** | **95.23% (.07)** |

**Discussion**

Here we tested the effects of cognitive load with either emotional or non-emotional properties on valence bias. As predicted, categorizations of surprise were more negative under emotional load than non-emotional load. This result extends previous work showing that higher non-emotional load did not affect categorizations (Mattek et al., 2016), and aligns with literature demonstrating that the emotional properties of cognitively demanding tasks impact both behavioral and neural responses during those tasks (Egner et al., 2008). We also found evidence that response trajectories were modulated by cognitive load as a function of trial-by-trial categorizations. Specifically, previous work has shown that positive categorizations of surprised faces are associated with greater response competition (i.e., more attraction to the competing – negative - response) than negative categorizations (Brown et al., 2017), and here we demonstrate that this difference is mitigated under high cognitive load. In other words, negative categorizations were associated with increased response competition under high load compared to low load, whereas response competition during positive categorizations tended to decrease, although this change did not reach statistical significance. This parallels other work showing that high non-emotional load increases distractor processing (Lavie & De Fockert, 2005) and response competition measured with mouse-based response trajectories (Bundt, Ruitenberg, Abrahamse, & Notebaert, 2018). We discuss these results in the context of the initial negativity hypothesis below.

**Domain-specific effects**

The initial negativity hypothesis posits that positive categorizations of ambiguous stimuli rely on regulatory resources that likely override an initial negativity (Neta et al., 2009; Petro et al., 2018). We used a standard working memory paradigm (Ahmed, 2018; Burnham, 2010; Lavie & De Fockert, 2005) to induce high cognitive load with either emotional or non-emotional properties while participants made valence judgments of surprised facial expressions. As expected, participants categorized surprise as more negative under emotional loads (i.e., when the emotional resources likely required for a positive interpretation are being depleted), most likely due to a reliance on overlapping domain-specific resources. These findings lend insight into the resources likely needed for positive categorizations, particularly in light of previous work demonstrating that non-emotional (numeric) load does not appear to affect categorizations of surprised faces (Mattek et al., 2016).

These findings are also consistent with work showing recruitment of dissociable neural resources for cognitively demanding tasks as a function of the load domain (Etkin et al., 2006; Neta et al., 2011). For instance, performing an emotional expression-based n-back task recruited greater amygdala activation when compared to an identity-based task (Neta & Whalen, 2011). Given the initial negativity hypothesis’ prediction that positivity relies on regulatory resources, it may be that emotional loads depleted these resources by recruiting regions that are functionally connected with the amygdala and are important for emotion regulation. One such region, the ventromedial prefrontal cortex (vmPFC) shows anatomical, functional, and structural connectivity with the amygdala (Amaral et al., 1992; Milad & Quirk, 2002; Johansen-Berg et al., 2008; Kim & Whalen, 2009; Amaral, 1992; Ghashghaei et al., 2007). and shows inverse activity patterns with amygdala (e.g., vmPFC increases as amygdala decreases) when subjects are asked to regulate their emotions (Ochsner et al., 2002; Jackson et al., 2003; Urry et al., van Reekum et al., 2007) and in response to positive categorizations of surprised faces (Kim et al., 2003).

Critically, these results also corroborate evidence that resource depletion is associated with a greater reliance on bottom-up emotional responses and weaker top-down responses. For example, some work has shown negative images recruited greater amygdala and less prefrontal activity under depletion (Wagner & Heatherton, 2013). In the context of surprised faces, the initial bottom-up response associated with greater amygdala activity results in more negative categorizations (Kim et al., 2003; Neta & Whalen, 2010; Petro et al., 2018). Thus, the increase in negative categorizations of surprise under emotional load supports the initial negativity hypothesis.

**Domain-general effects**

While subjective categorizations of ambiguity were susceptible to the load domain, the response competition was instead vulnerable to domain-general cognitive demands. That is, maximum deviations varied as a function of low compared to high load, regardless of the load domain. Specifically, under a low load, positive categorizations are typically associated with greater response competition than negative categorizations, but this effect was no longer present under high load. Further, this effect was driven by an *increase* in response competition for negative trials under high load, although there was also a non-significant trend for response competition for positive trials to *decrease* under load. This replicates previous work showing that the response competition underlying the valence bias task is susceptible to increased cognitive demands generally (Mattek et al., 2016). One interpretation of these findings is that response competition may reflect a type of distraction effect (Spivey, Grosjean, & Knoblich, 2005), consistent with effects showing high load results in deficits in inhibiting task-irrelevant information (Lavie, Hirst, de Fockert, & Vidling, 2004).

Previous work has shown that emotional ambiguity categorization relies on a domain-general task control network called the cingulo-opercular network (Neta et al., 2013); though speculative, the cognitive loads may have taxed these resources, as this network is recruited in response to many types of ambiguity (Neta et al., 2013; Neta et al., 2014; Sterzer, Russ, Preibisch, & Kleinschmidt, 2002; Thompson-Schill, D’Esposito, Aguirre, & Farah, 1997). Other neuroimaging work supports the notion that cognitive loads would preoccupy resources in the cingulo-opercular network; for instance, regions in the network (i.e., anterior cingulate cortex and anterior insula) regularly show activity increases during cognitively demanding tasks, such as those requiring increased attention and control (Duncan & Owen, 2000; Nee, Wager, & Jonides, 2007). As such, the demands induced during high cognitive load, regardless of the emotional properties of the load, likely increased demands in this network. Ultimately, this increase in demands for this domain-general network are one explanation for the observed increase in response competition (i.e., maximum deviations) for negative ratings during high cognitive load.

**Limitations and future directions**

There are a few limitations to the present study. First, acuracy on the memory probe task, even under high load, was high, suggesting that the cognitive resources were not depleted heavily. Relatedly, participants may have been able to rely on recognition (rather than active working memory maintenance) for the memory probes, as humans are readily able to identify previously seen images after exposure to a large amount of material (i.e., 600 images) at high accuracy (Shepard, 1967). Relatedly, in the present study, each image appeared within only one image matrix and each matrix was only presented once, perhaps facilitating participants’ ability to recognize the image during the memory probe. Future work could address this by increasing the demands of the task, either by using more than six images in the high load matrix, re-using some images across trials making it more difficult to remember in the image probe was presented on that specific trial, or making the probe task more difficult (e.g., testing the location of the image in the previous matrix rather than just a present/not judgment).

Further, we attempted to use a similar working memory task that could directly compare emotional versus non-emotional properties. Thus, unlike previous work that used numerical sequences that could be rehearsed using verbal working memory, our task likely relies more on visual working memory (Baddeley, 1998). In the context of visual working memory, perhaps one interesting avenue for future work is to incorporate eye tracking to explore which images participants attended to the most within a matrix, offering insight into which images may be most likely to be held in working memory. In turn, this would allow testing on a trial-by-trial basis, such that attention towards either positive or negative emotional images could be quantified and explored in the context of subsequent ratings of surprised expressions.

Finally, the effects of high load on response trajectories are different from those of stress; that is, negative categorizations of surprise are associated with greater response competition under high load but lower response competition under stress. It is possible that different resources are depleted in these different circumstances, and/or that a different network of brain regions are required for task performance when under high load or high stress. Future work will be needed to disentangle these effects.

**Conclusions**

Here we have provided both a conceptual replication and a novel extension of previous work which tested the effects of cognitive load on categorizations of ambiguity (Mattek et al., 2016). Notably, these findings illuminate the processes putatively needed for arriving at a positive categorization by demonstrating that, under emotional load, these positive categorizations were less likely. As such, these findings lend further support for the initial negativity hypothesis by suggesting that positivity (more so than negativity) relies on additional emotion-related resources. We also demonstrated a domain-general effect of load on response competition, which is likely related to the domain-general demands of high load within the cingulo-opercular network. Future work should explore the underlying neural mechanisms of these processes. Notably, elucidating the neural mechanisms through which individuals become more negative would offer insight into a range of clinical disorders characterized by negativity bias (e.g., anxiety, depression). Further, this work may even shed light on mechanisms through which those in cognitively and emotionally demanding positions (e.g., healthcare workers) experience negativity related to workplace burnout.

**Acknowlegements:** We thank Hannah E. Raila for early discussions about the hypotheses and approach. We thank Michael D. Dodd and Jeffrey R. Stevens for suggestions and comments on an earlier version of the manuscript. We thank Rebecca L. Brock for statistical consultation. And we thank Catherine C. Brown, Nathan M. Petro, and Kayla Clark for feedback on a later draft of the manuscript.

# References

Ahmed, L. (2018). Knowing how you are feeling depends on what’s on my mind: Cognitive load and expression categorization. *Emotion*, *18*(2), 190–201. doi:[10.1037/emo0000312](https://doi.org/10.1037/emo0000312)

Baddeley, A. D. (1986). Working memory. *Philosophical Transactions of the Royal Society of London*, *302*(110), 311–324.

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest: A Journal of the American Psychological Society*, *20*(1), 1–68. doi:[10.1177/1529100619832930](https://doi.org/10.1177/1529100619832930)

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1-48. doi:10.18637/jss.v067.i01.

Baumeister, R. F., & Heatherton, T. F. (1996). Self-regulation failure: An overview. *Psychological Inquiry*, *7*(1), 1–15. doi:[10.1207/s15327965pli0701\_1](https://doi.org/10.1207/s15327965pli0701_1)

Blair, K. S., Smith, B. W., Mitchell, D. G. V., Morton, J., Vythilingam, M., Pessoa, L., … Blair, R. J. R. (2007). Modulation of emotion by cognition and cognition by emotion. *NeuroImage*, *35*(1), 430–440. doi:[10.1016/j.neuroimage.2006.11.048](https://doi.org/10.1016/j.neuroimage.2006.11.048)

Brown, C. C., Raio, C. M., & Neta, M. (2017). Cortisol responses enhance negative valence perception for ambiguous facial expressions. *Scientific Reports*, *7*(1), 15107. doi:[10.1038/s41598-017-14846-3](https://doi.org/10.1038/s41598-017-14846-3)

Bundt, C., Ruitenberg, M. F. L., Abrahamse, E. L., & Notebaert, W. (2018). Early and late indications of item-specific control in a stroop mouse tracking study. *PLOS ONE*, *13*(5), e0197278. doi:[10.1371/journal.pone.0197278](https://doi.org/10.1371/journal.pone.0197278)

Burnham, B. R. (2010). Cognitive load modulates attentional capture by color singletons during effortful visual search. *Acta Psychologica*, *135*(1), 50–58. doi:[10.1016/j.actpsy.2010.05.003](https://doi.org/10.1016/j.actpsy.2010.05.003)

Calcagnì, A., Lombardi, L., & Sulpizio, S. (2017). Analyzing spatial data from mouse tracker methodology: An entropic approach. *Behavior Research Methods*, *49*(6), 2012–2030. doi:[10.3758/s13428-016-0839-5](https://doi.org/10.3758/s13428-016-0839-5)

Carroll, J. M., & Russell, J. A. (1996). Do facial expressions signal specific emotions? Judging emotion from the face in context. *Journal of Personality and Social Psychology*, *70*(2), 205–218. doi:[10.1037//0022-3514.70.2.205](https://doi.org/10.1037//0022-3514.70.2.205)

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, *8*(4), 293–332. doi:[10.1207/s1532690xci0804\_2](https://doi.org/10.1207/s1532690xci0804_2)

Darwin, C. (1872). *The expression of the emotions in man and animals*. John Murray.

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, *23*(10), 475–483. doi:[10.1016/s0166-2236(00)01633-7](https://doi.org/10.1016/s0166-2236(00)01633-7)

Egner, T., Etkin, A., Gale, S., & Hirsch, J. (2008). Dissociable neural systems resolve conflict from emotional versus nonemotional distracters. *Cerebral Cortex (New York, N.Y.: 1991)*, *18*(6), 1475–1484. doi:[10.1093/cercor/bhm179](https://doi.org/10.1093/cercor/bhm179)

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, *17*(2), 124–129. doi:[10.1037/h0030377](https://doi.org/10.1037/h0030377)

Etkin, A., Egner, T., Peraza, D. M., Kandel, E. R., & Hirsch, J. (2006). Resolving emotional conflict: A role for the rostral anterior cingulate cortex in modulating activity in the amygdala. *Neuron*, *51*(6), 871–882. doi:[10.1016/j.neuron.2006.07.029](https://doi.org/10.1016/j.neuron.2006.07.029)

Flexas, A., Rosselló, J., Christensen, J. F., Nadal, M., Rosa, A. O. L., & Munar, E. (2013). Affective priming using facial expressions modulates liking for abstract art. *PLOS ONE*, *8*(11), e80154. doi:[10.1371/journal.pone.0080154](https://doi.org/10.1371/journal.pone.0080154)

Freeman, J. B., & Ambady, N. (2010). MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method. *Behavior Research Methods*, *42*(1), 226–241. doi:[10.3758/BRM.42.1.226](https://doi.org/10.3758/BRM.42.1.226)

Freeman, J., Dale, R., & Farmer, T. (2011). Hand in motion reveals mind in motion. *Frontiers in Psychology*, *2*. doi:[10.3389/fpsyg.2011.00059](https://doi.org/10.3389/fpsyg.2011.00059)

Frijda, N. H. (1986). *The emotions*. Paris, France: Editions de la Maison des Sciences de l’Homme.

Frith, C. (2009). Role of facial expressions in social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3453–3458. doi:[10.1098/rstb.2009.0142](https://doi.org/10.1098/rstb.2009.0142)

Green, C., & Guo, K. (2018). Factors contributing to individual differences in facial expression categorisation. *Cognition & Emotion*, *32*(1), 37–48. doi:[10.1080/02699931.2016.1273200](https://doi.org/10.1080/02699931.2016.1273200)

Hehman, E., Stolier, R. M., & Freeman, J. B. (2015). Advanced mouse-tracking analytic techniques for enhancing psychological science. *Group Processes & Intergroup Relations*, *18*(3), 384–401. doi:[10.1177/1368430214538325](https://doi.org/10.1177/1368430214538325)

Izard, C. E. (1994). Innate and universal facial expressions: Evidence from developmental and cross-cultural research. *Psychological Bulletin*, *115*(2), 288–299. doi:[10.1037/0033-2909.115.2.288](https://doi.org/10.1037/0033-2909.115.2.288)

Jiaping, C., Yuejia, L. U. O., Fang, C. U. I., Jiaping, C., Yuejia, L. U. O., & Fang, C. U. I. (2017). Empathy for pain influenced by cognitive load: Evidence from an ERP study. *Acta Psychologica Sinica*, *49*(5), 622–630. doi:[10.3724/SP.J.1041.2017.00622](https://doi.org/10.3724/SP.J.1041.2017.00622)

Kieslich, P.J. , Henninger, F., Wulff, D. U., Haslbeck, J. M., Schulte-Mecklenbeck, M. (2019). Mouse-tracking: A practical guide to implementation and analysis. In Schulte-Mecklenbeck M, Kühberger A, Johnson JG (eds.), *A Handbook of Process Tracing Methods*, 111-130. Routledge, New York, NY.

Kim, H., Somerville, L. H., Johnstone, T., Alexander, A. L., & Whalen, P. J. (2003). Inverse amygdala and medial prefrontal cortex responses to surprised faces. *Neuroreport*, *14*(18), 2317–2322. doi:[10.1097/00001756-200312190-00006](https://doi.org/10.1097/00001756-200312190-00006)

Kim, H., Somerville, L. H., Johnstone, T., Polis, S., Alexander, A. L., Shin, L. M., & Whalen, P. J. (2004). Contextual modulation of amygdala responsivity to surprised faces. *Journal of Cognitive Neuroscience*, *16*(10), 1730–1745. doi:[10.1162/0898929042947865](https://doi.org/10.1162/0898929042947865)

Knight, M., Seymour, T. L., Gaunt, J. T., Baker, C., Nesmith, K., & Mather, M. (2007). Aging and goal-directed emotional attention: Distraction reverses emotional biases. *Emotion*, *7*(4), 705–714. doi:[10.1037/1528-3542.7.4.705](https://doi.org/10.1037/1528-3542.7.4.705)

Krieglmeyer, R., Deutsch, R., De Houwer, J., & De Raedt, R. (2010). Being moved: Valence activates approach-avoidance behavior independently of evaluation and approach-avoidance intentions. *Psychological Science*, *21*(4), 607–613. doi:[10.1177/0956797610365131](https://doi.org/10.1177/0956797610365131)

Kron, A., Schul, Y., Cohen, A., & Hassin, R. R. (2010). Feelings don’t come easy: Studies on the effortful nature of feelings. *Journal of Experimental Psychology: General*, *139*(3), 520–534. doi:[10.1037/a0020008](https://doi.org/10.1037/a0020008)

Kujawa, A., Wu, M., Klumpp, H., Pine, D. S., Swain, J. E., Fitzgerald, K. D., … Phan, K. L. (2016). Altered development of amygdala-anterior cingulate cortex connectivity in anxious youth and young adults. *Biological Psychiatry : Cognitive Neuroscience and Neuroimaging*, *1*(4), 345–352. doi:[10.1016/j.bpsc.2016.01.006](https://doi.org/10.1016/j.bpsc.2016.01.006)

Lang, P., Bradley, M. M., & Cuthbert, B. N. (2008). International affective picture system (IAPS): Affective ratings of pictures and instruction manual., Technical Report A–8. University of Florida, Gainesville, FL.

Lavie, N., & De Fockert, J. (2005). The role of working memory in attentional capture. *Psychonomic Bulletin & Review*, *12*(4), 669–674. doi:[10.3758/BF03196756](https://doi.org/10.3758/BF03196756)

Lavie, N., Hirst, A., Fockert, J. W. de, & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology. General*, *133*(3), 339–354. doi:[10.1037/0096-3445.133.3.339](https://doi.org/10.1037/0096-3445.133.3.339)

Lazarus, R. S., & Alfert, E. (1964). Short-circuiting of threat by experimentally altering cognitive appraisal. *The Journal of Abnormal and Social Psychology*, *69*(2), 195–205. doi:[10.1037/h0044635](https://doi.org/10.1037/h0044635)

Lundqvist, D., Flykt, A., & Öhman, A. (1998). The karolinska directed emotional faces—KDEF (CD ROM)., Stockholm: Karolinska Institute, Departmentof Clinical Neuroscience, PsychologySection.

Mather, M., & Carstensen, L. L. (2005). Aging and motivated cognition: The positivity effect in attention and memory. *Trends in Cognitive Sciences*, *9*(10), 496–502. doi:[10.1016/j.tics.2005.08.005](https://doi.org/10.1016/j.tics.2005.08.005)

Mattek, A. M., Whalen, P. J., Berkowitz, J. L., & Freeman, J. B. (2016). Differential effects of cognitive load on subjective versus motor responses to ambiguously valenced facial expressions. *Emotion*, *16*(6), 929–936. doi:[10.1037/emo0000148](https://doi.org/10.1037/emo0000148)

Muraven, M., Tice, D. M., & Baumeister, R. F. (1998). Self-control as a limited resource: Regulatory depletion patterns. *Journal of Personality and Social Psychology*, *74*(3), 774–789. doi:[10.1037/0022-3514.74.3.774](https://doi.org/10.1037/0022-3514.74.3.774)

Murphy, G., Groeger, J. A., & Greene, C. M. (2016). Twenty years of load theory—where are we now, and where should we go next? *Psychonomic Bulletin & Review*, *23*(5), 1316–1340. doi:[10.3758/s13423-015-0982-5](https://doi.org/10.3758/s13423-015-0982-5)

Nagamatsu, L. S., Voss, M., Neider, M. B., Gaspar, J. G., Handy, T. C., Kramer, A. F., & Liu-Ambrose, T. Y. L. (2011). Increased cognitive load leads to impaired mobility decisions in seniors at risk for falls. *Psychology and Aging*, *26*(2), 253–259. doi:[10.1037/a0022929](https://doi.org/10.1037/a0022929)

Nee, D. E., Wager, T. D., & Jonides, J. (2007). Interference resolution: Insights from a meta-analysis of neuroimaging tasks. *Cognitive, Affective, & Behavioral Neuroscience*, *7*(1), 1–17. doi:[10.3758/CABN.7.1.1](https://doi.org/10.3758/CABN.7.1.1)

Neta, M., Davis, F. C., & Whalen, P. J. (2011). Valence resolution of ambiguous facial expressions using an emotional oddball task. *Emotion*, *11*(6), 1425–1433. doi:[10.1037/a0022993](https://doi.org/10.1037/a0022993)

Neta, M., Kelley, W. M., & Whalen, P. J. (2013). Neural responses to ambiguity involve domain-general and domain-specific emotion processing systems. *Journal of Cognitive Neuroscience*, *25*(4), 547–557. doi:[10.1162/jocn\_a\_00363](https://doi.org/10.1162/jocn_a_00363)

Neta, M., Norris, C. J., & Whalen, P. J. (2009). Corrugator muscle responses are associated with individual differences in positivity-negativity bias. *Emotion (Washington, D.C.)*, *9*(5), 640–648. doi:[10.1037/a0016819](https://doi.org/10.1037/a0016819)

Neta, M., Schlaggar, B. L., & Petersen, S. E. (2014). Separable responses to error, ambiguity, and reaction time in cingulo-opercular task control regions. *NeuroImage*, *99*, 59–68. doi:[10.1016/j.neuroimage.2014.05.053](https://doi.org/10.1016/j.neuroimage.2014.05.053)

Neta, M., & Tong, T. T. (2016). Don’t like what you see? Give it time: Longer reaction times associated with increased positive affect. *Emotion*, *16*(5), 730–739. doi:[10.1037/emo0000181](https://doi.org/10.1037/emo0000181)

Neta, M., & Whalen, P. J. (2010). The primacy of negative interpretations when resolving the valence of ambiguous facial expressions. *Psychological Science*, *21*(7), 901–907. doi:[10.1177/0956797610373934](https://doi.org/10.1177/0956797610373934)

Petro, N. M., Tong, T. T., Henley, D. J., & Neta, M. (2018). Individual differences in valence bias: fMRI evidence of the initial negativity hypothesis. *Social Cognitive and Affective Neuroscience*, *13*(7), 687–698. doi:[10.1093/scan/nsy049](https://doi.org/10.1093/scan/nsy049)

Pontari, B. A., & Schlenker, B. R. (2000). The influence of cognitive load on self-presentation: Can cognitive busyness help as well as harm social performance? *Journal of Personality and Social Psychology*, *78*(6), 1092–1108. doi:[10.1037/0022-3514.78.6.1092](https://doi.org/10.1037/0022-3514.78.6.1092)

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

Said, C. P., & Todorov, A. (2011). A statistical model of facial attractiveness. *Psychological Science*, *22*(9), 1183–1190. doi:[10.1177/0956797611419169](https://doi.org/10.1177/0956797611419169)

Scalf, P. E., Torralbo, A., Tapia, E., & Beck, D. M. (2013). Competition explains limited attention and perceptual resources: Implications for perceptual load and dilution theories. *Frontiers in Psychology*, *4*. doi:[10.3389/fpsyg.2013.00243](https://doi.org/10.3389/fpsyg.2013.00243)

Sheppes, G., & Meiran, N. (2008). Divergent cognitive costs for online forms of reappraisal and distraction. *Emotion*, *8*(6), 870–874. doi:[10.1037/a0013711](https://doi.org/10.1037/a0013711)

Sterzer, P., Russ, M. O., Preibisch, C., & Kleinschmidt, A. (2002). Neural correlates of spontaneous direction reversals in ambiguous apparent visual motion. *NeuroImage*, *15*(4), 908–916. doi:[10.1006/nimg.2001.1030](https://doi.org/10.1006/nimg.2001.1030)

Storbeck, J. (2012). Performance costs when emotion tunes inappropriate cognitive abilities: Implications for mental resources and behavior. *Journal of Experimental Psychology: General*, *141*(3), 411–416. doi:[10.1037/a0026322](https://doi.org/10.1037/a0026322)

Thomas, L., Donohue-Porter, P., & Stein Fishbein, J. (2017). Impact of interruptions, distractions, and cognitive load on procedure failures and medication administration errors: *Journal of Nursing Care Quality*, *32*(4), 309–317. doi:[10.1097/NCQ.0000000000000256](https://doi.org/10.1097/NCQ.0000000000000256)

Thompson-Schill, S. L., D’Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(26), 14792–14797. doi:[10.1073/pnas.94.26.14792](https://doi.org/10.1073/pnas.94.26.14792)

Todorov, A., Baron, S. G., & Oosterhof, N. N. (2008). Evaluating face trustworthiness: A model based approach. *Social Cognitive and Affective Neuroscience*, *3*(2), 119–127. doi:[10.1093/scan/nsn009](https://doi.org/10.1093/scan/nsn009)

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., … Nelson, C. (2009a). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242–249. doi:[10.1016/j.psychres.2008.05.006](https://doi.org/10.1016/j.psychres.2008.05.006)

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., … Nelson, C. (2009b). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242–249. doi:[10.1016/j.psychres.2008.05.006](https://doi.org/10.1016/j.psychres.2008.05.006)

Trémolière, B., Gagnon, M.-È., & Blanchette, I. (2016). Cognitive load mediates the effect of emotion on analytical thinking. *Experimental Psychology*, *63*(6), 343–350. doi:[10.1027/1618-3169/a000333](https://doi.org/10.1027/1618-3169/a000333)

Van Dillen, L. F., Heslenfeld, D. J., & Koole, S. L. (2009). Tuning down the emotional brain: An fMRI study of the effects of cognitive load on the processing of affective images. *NeuroImage*, *45*(4), 1212–1219. doi:[10.1016/j.neuroimage.2009.01.016](https://doi.org/10.1016/j.neuroimage.2009.01.016)

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag: New York.