Caffeine cravings impair memory and metacognition

Matthew A. Palmer, James D. Sauer, Angus Ling & Joshua Riza

To cite this article: Matthew A. Palmer, James D. Sauer, Angus Ling & Joshua Riza (2017) Caffeine cravings impair memory and metacognition, Memory, 25:9, 1225-1234, DOI: 10.1080/09658211.2017.1282968

To link to this article: http://dx.doi.org/10.1080/09658211.2017.1282968

Published online: 16 Feb 2017.
Caffeine cravings impair memory and metacognition

Matthew A. Palmer\textsuperscript{a}, James D. Sauer\textsuperscript{b}, Angus Ling\textsuperscript{b} and Joshua Riza\textsuperscript{b}

\textsuperscript{a}Division of Psychology, School of Medicine, University of Tasmania, Launceston, Tasmania, Australia; \textsuperscript{b}Division of Psychology, School of Medicine, University of Tasmania, Hobart, Tasmania, Australia

\begin{abstract}
Cravings for food and other substances can impair cognition. We extended previous research by testing the effects of caffeine cravings on cued-recall and recognition memory tasks, and on the accuracy of judgements of learning (JOLs; predicted future recall) and feeling-of-knowing (FOK; predicted future recognition for items that cannot be recalled). Participants ($N = 55$) studied word pairs (POND-BOOK) and completed a cued-recall test and a recognition test. Participants made JOLs prior to the cued-recall test and FOK judgements prior to the recognition test. Participants were randomly allocated to a craving or control condition; we manipulated caffeine cravings via a combination of abstinence, cue exposure, and imagery. Cravings impaired memory performance on the cued-recall and recognition tasks. Cravings also impaired resolution (the ability to distinguish items that would be remembered from those that would not) for FOK judgements but not JOLs, and reduced calibration (correspondence between predicted and actual accuracy) for JOLs but not FOK judgements. Additional analysis of the cued-recall data suggested that cravings also reduced participants’ ability to monitor the likely accuracy of answers during the cued-recall test. These findings add to prior research demonstrating that memory strength manipulations have systematically different effects on different types of metacognitive judgements.
\end{abstract}

Cravings are a motivational state characterised by an intense desire for a particular substance (Kemps & Tiggemann, 2010). Cravings are experienced by a large percentage of the population and can occur for a variety of illegal and legal substances, such as caffeine, chocolate, nicotine, and cocaine (e.g., Juliano & Griffiths, 2004; Kemps & Tiggemann, 2009; Kilts et al., 2001; Lafay et al., 2001). The negative impacts of cravings include increased consumption of unhealthy foods and illicit substances (e.g., McManus & Waller, 1995; Tiffany, 1990). Perhaps less intuitively, cravings have also been shown to impair cognitive processes such as attention, executive functioning, language comprehension, and working memory (e.g., Baxter & Hinson, 2001; Green, Rogers, & Elliman, 2000; Kemps, Tiggemann, & Grigg, 2008; Uva et al., 2010; Zwaan, Stanfield, & Madden, 2000).

Tiffany’s (1990) prominent model can be used to understand why cravings impair cognition. Consumption behaviours, such as drinking coffee and eating chocolate, are stored in long-term memory as action schemata. If regularly performed, these behaviours become increasingly automated and efficient, and difficult to inhibit if triggered by cues (e.g., the sight of chocolate or smell of coffee). Inhibition of the action schema does not occur automatically; it requires cognitive processing and attention. Because cognitive resources are finite, the inhibition of the consumption schema leaves fewer cognitive resources for other tasks. Hence, cravings impair performance on concurrent tasks that require cognitive resources (Kemps et al., 2008; Tiffany & Conklin, 2000).

The present research extends understanding of the effects of cravings on cognition in two ways. First, we examine the effects of cravings on recall-based and recognition-based episodic memory. Tiffany’s (1990) model predicts that cravings will impair performance on memory tasks, given that these require cognitive resources. Recent research has found that cravings impaired performance on recall-based memory tasks but had minimal effect on recognition-based tasks (Zuj, Palmer, & Kemps, 2015). Given that recognition-based tasks require fewer cognitive resources than recall-based tasks (e.g., Anderson, Craik, & Naveh-Benjamin, 1998) these findings raise the possibility that memory tasks that demand fewer resources may not be susceptible to the effects of cravings. We test whether these results can be replicated.

Second, we examine the effects of cravings on the accuracy of metacognitive judgements. Metacognition refers to an individual’s knowledge of their own cognitive processes, and ability to monitor and predict memory performance (e.g., Flavell, 1979; Hart, 1965). Accurate metacognitive judgements are crucial for memory and decision-making in many applied settings. For example,
students studying for exams must make judgements about how well they have learned course content (e.g., Metcalfe & Finn, 2008). People can generally predict future memory performance reasonably well (e.g., Nelson & Dunlosky, 1991; Vesonder & Voss, 1985). However, factors that affect memory performance can also affect the accuracy of metacognitive judgements (e.g., Bothwell, Deffenbacher, & Brigham, 1987; Deffenbacher, 1980; Hertzog, Dunlosky, & Sinclair, 2010; Palmer, Brewer, Weber, & Nagesh, 2013; Souchay & Isingrini, 2012), and cravings could reduce the accuracy of metacognitive judgements in several ways.

Cravings could also inflate overconfidence about the likelihood of remembering information in the future. Typically, overconfidence in metacognitive judgements increases as memory tasks become more difficult (the hard-easy effect; Gigerenzer, Hoffrage, & Kleinbölting, 1991). Accordingly, many factors that reduce memory performance produce overestimation in memory predictions. For example, increases in retention interval impair memory but have little impact on predictions of future remembering (Koriat, Bjork, Sheffer, & Bar, 2004). To the extent that cravings impair memory, they might also inflate overconfidence in future memory performance. Increased overconfidence may, for example, cause students to overestimate how well they have learned material and underestimate how much more study is required (e.g., Metcalfe & Finn, 2008; Nelson & Dunlosky, 1991).

Cravings could also impair an individual’s ability to distinguish material that will be remembered from material that will not be remembered. This property of metacognition has been termed resolution, or discrimination (Murphy, 1973; Yaniv, Yates, & Smith, 1991). Impaired resolution would, for example, hamper students’ decisions about which material to allocate study time to. There is evidence that strength of memory can affect resolution, such that weaker memory is associated with poorer resolution (e.g., Bothwell et al., 1987; Perfect & Stollery, 1993), although this is not always the case (e.g., Palmer et al., 2013; Sauer, Brewer, Zweck, & Weber, 2010). Broadly, the rationale for this relationship is that the weaker the memory, the less memorial information will be accessible to evaluate during metacognitive judgements, and the less able one will be to distinguish material that will be remembered from material that will not. Specific versions of this account have been termed the memory constraint hypothesis (Hertzog et al., 2010) and the optimality hypothesis (Bothwell et al., 1987; Deffenbacher, 1980).

Importantly, the relationship between memory strength and resolution depends on the type of metacognitive judgement. Common measures of metacognition include judgements of learning (JOLs; Nelson & Dunlosky, 1991), which are predictions about the likely recall of studied material on a future test, and feeling-of-knowing (FOK), which refers to a subjective experience of feeling that one knows information despite being currently unable to recall it. FOK is often measured by asking for predictions of the likelihood of recognising a studied but currently inaccessible item on a future multiple-choice test (Hart, 1965). Note that JOLs are made prior to a memory test and concern future recall of target items, whereas FOK judgements are made at the test stage (after recall has failed) and concern future recognition of target items.

Evidence from research on ageing and episodic memory shows that deficits in memory strength in older adults – compared to younger adults – are associated with impaired resolution for FOK judgements but not JOLs (e.g., Hertzog et al., 2010; Hertzog & Hultsch, 2000; Souchay, Moulin, Clarys, Taconnat, & Isingrini, 2007; Thomas, Bulevich, & Dubois, 2011). This is because accurate JOLs and FOK judgements rely to some extent on different memorial information. The resolution of FOK judgements – made without retrieval of the target – is thought to rely on the retrieval of accurate partial information (e.g., some elements of the target) or contextual information related to the non-retrieved target item (Hertzog et al., 2010; Isingrini et al., 2016; Koriat, 1993; Sacher, Isingrini, & Taconnat, 2013; Souchay & Isingrini, 2012). The more accurate partial or contextual information retrieved, the better able a person is to judge their future recognition; hence, the more accurate the FOK judgement. Because the retrieval of partial and contextual information retrieves, the better able a person is to judge their future recognition; hence, the more accurate the FOK judgement. Because the retrieval of partial and contextual information varies with memory strength (e.g., Koriat, 1993; Yonelinas, 2002), FOK resolution decreases as memory strength decreases (Hertzog et al., 2010).

In contrast, JOLs can be based on information such as fluency of encoding (when JOLs are made immediately after study) or fluency of retrieval of the target (when made after a delay; Nelson & Dunlosky, 1991). As a result, JOL resolution is not necessarily dependent on memory strength. For example, when making delayed JOLs under strong memory conditions, a person might fluently retrieve many items. These items will be given high JOLs and have a high likelihood of being successfully remembered on a later test (e.g., Spellman & Bjork, 1992). Items that are not fluently retrieved will, on average, be given lower JOLs and be less likely to be remembered on a later test. When making JOLs under weak memory conditions, fluent retrieval will likely occur for fewer items. However, any items that are fluently retrieved will still be given high JOLs and will likely be remembered on a later test, and items that are not fluently retrieved will be given low JOLs and will be less likely to be remembered later. Thus, JOL resolution can remain stable across conditions that differ in memory strength because when making JOLs, failure to retrieve a target is diagnostic of future failure to remember the target on a future test.

It is an empirical question whether cravings will produce similar effects on JOLs and FOK judgements to those produced by differences in age. However, there is precedent that FOK resolution can be influenced by experimentally manipulated factors that affect memory strength, such as divided attention and repeated presentation of items (Hertzog et al., 2010; Sacher et al., 2013; Sacher, Taconnat,
Souchay, & Isingrini, 2009). To the extent that these effects generalise to cravings, resolution will be reduced for FOK judgements but not JOLs.

To investigate the effects of cravings on memory performance and metacognition, we used cravings for coffee. In Western culture, coffee is the most common source of caffeine, a commonly craved substance (Frary, Johnson, & Wang, 2005; Juliano & Griffiths, 2004; Kemps & Tiggemann, 2009). Although caffeine cannot be the basis for a diagnosis of a substance abuse disorder according to the Diagnostic and Statistical Manual of Mental Disorders (5th edition; DSM-5; American Psychiatric Association, 2013), there are many similarities between the effects of caffeine cravings and the effects of cravings for substances that form the basis for substance use disorders and addiction. For example, among heavy consumers of caffeine (i.e., people who consume 350 mg or more of caffeine per day, on average), cravings are positively correlated with attentional bias toward caffeine-related stimuli (Stafford & Yeomans, 2005; Yeomans, Javerian, Tovey, & Stafford, 2005). Similar relationships between cravings and attentional bias have been found among users of cocaine and heroin (e.g., Franken, Kroon, Wiers, & Jansen, 2000; Garavan et al., 2000). Thus, although our research dealt only with caffeine cravings, there is reason to expect that the results may generalise to other craved substances.

Following prior research (e.g., Kemps et al., 2008), we manipulated coffee cravings using a combination of abstinence, imagery, and in-vivo exposure to consumption cues. Participants were allocated to the craving or control condition (between-subjects) and studied a series of word pairs, then completed cued-recall and recognition memory tasks. Participants predicted their memory performance via JOLs (prior to the cued-recall test) and FOK judgements (prior to the recognition test).

We predicted that cravings would reduce memory performance on the cued-recall task. We did not have clear expectations about the effects of cravings on performance on the 4-alternative forced choice (4AFC) recognition task because prior research suggests that cravings may not impair recognition memory (Zuj et al., 2015). We also predicted that cravings would reduce resolution for FOK judgements but not JOLs (drawing on previous research by Hertzog et al., 2010; Sacher et al., 2009; 2013), and we tested whether cravings affected the correspondence between metacognitive judgements and objective accuracy on the memory tests (e.g., by producing hard-easy effects; Gigerenzer et al., 1991).

**Method**

**Participants and screening procedures**

Participants were 55 coffee drinkers recruited from the University of Tasmania, Flinders University, and the broader community (36 females, 17 males, 2 reported sex as “other”; mean age = 29.9 years, SD = 12.2). Participants received course credit or a $20 gift voucher. The recruiting advertisements for this study stated that the study required coffee drinkers, and all participants reported drinking at least one coffee per day on average.

Participants completed a version of the Leeds Dependence Questionnaire (Raistrick et al., 1994) adapted for use with coffee. Questions asked participants to report the frequency with which they experienced 10 characteristics of dependence during the previous week (e.g., “Do you find yourself thinking about when you will be able to have another coffee?”) with responses made on a 4-point scale (0 = never, 1 = sometimes, 2 = often, 3 = always). Eleven participants had an average response on this scale of below one (sometimes). These participants completed the study but were excluded from analyses on the basis that the craving manipulation would not be relevant for people who had little potential to experience cravings. Thus, 66 people in total completed the study and 55 were included in the analyses (those with an average score of ≥1 on the adapted LDQ).

**Materials and procedure**

Participants were randomly allocated to the craving (n = 27) and control groups (n = 28) and completed the experiment in a laboratory individually or in small groups (maximum of three participants). All participants completed demographic questions, followed by a study phase, JOLs, a cued-recall test, FOK judgements, a recognition test, and manipulation checks.

**Stimuli**

The MRC Psycholinguistic Database (Coltheart, 1981) was used to create 100 sets of semantically unrelated concrete nouns that varied in the parameters of imagability (200–600), concreteness (300–700), and familiarity (100–700). This variation was designed to promote variation in memory performance across the stimulus set. Each set contained four words. Two words from each set were used in the JOL and cued-recall tasks (the same two from each set for every participant). All four words from each set were used in the FOK and recognition tests. Words were presented in 36-point capital letters, Times New Roman font against a white background, with a resolution of 1024 x 768, using E-Prime (Psychology Software Tools, 2012).

**Study phase**

Prior to the study phase, participants were informed that they would see a series of word pairs and their memory for the word pairs would be tested later. Participants then studied 100 cue-target word pairs (e.g., POND-BOOK) one-at-a-time. Order of presentation was randomised. Progress was self-paced with an enforced 30-second break after 50 word pairs. A break after 50 trials also occurred after 50 trials in the JOL, cued-recall, FOK, and recognition phases.
JOLs
After completing the study phase, participants made delayed JOLs. Participants viewed the first word of each pair (e.g., POND___?) and rated the likelihood they would remember the second word if shown the first word in a subsequent test trial (from 0%–certain that you will not remember the second word to 100%–certain that you will remember the second word). On each trial, participants were presented with a single cue word and progressed to the next cue word after making their response. Order of presentation was randomised with the constraint that words in the first 50 pairs for the study phase appeared in the first 50 for the JOL phase. This also applied to the cued-recall, FOK, and recognition phases.

Cued-recall test
Participants were then shown the first word of each pair (e.g., POND___?) and asked to type in the matching second word. Participants were told that if they did not know the answer it was preferable to leave the space blank. Cue words were presented one-at-a-time.

FOK judgements
Participants then rated the likelihood that they would be able to recognise the second word of each pair from a list of four words if they were shown the first word (e.g., POND____?). Participants were given an example trial for practice. Ratings were made on a scale from 0 –definitely not to 100 –definitely yes. Participants made FOK ratings for all word pairs (not only trials where they did not make a cued-recall response) but our analyses focused on trials for which participants did not make any recall response. Analyses including FOK responses to all pairs produced very similar patterns of results.

Recognition test
Participants were then presented with the first word of each pair along with four response options presented in a vertical list (e.g., POND____KILT; SALT; BOOK; FOWL). Participants were asked to choose the corresponding second word from the list of four options. Thus, this task was a 4AFC recognition decision. Trials in which participants did not choose a response option were coded as incorrect response. On all trials, the correct answer appeared among the alternatives. The placement of the correct answer was counterbalanced so that the correct answer appeared equally often in each of the four locations.

Craving manipulation
Following prior research (e.g., Kemps et al., 2008), we manipulated coffee cravings via a combination of abstinence, imagery, and in-vivo exposure to consumption cues. Participants in the craving condition were required to abstain from consuming coffee on the day of testing. Those in the control condition were asked to continue consuming coffee as normal on the day of testing. All participants reported that they had complied with these instructions.

All participants completed the imagery and exposure tasks twice during the experiment: between the study phase and delayed JOLs, and between the cued-recall task and FOK judgements. Participants in the craving group were presented with a jug of freshly brewed black coffee in a plunger and were told that under no circumstances were they allowed to consume the beverage. The experimenter then read aloud scripted instructions adopted from Green et al. (2000) to prompt interaction with the craved substance (e.g., “Pick up the jug of freshly made coffee and pour yourself a cup”), and mental imagery with strong sensory components (“Pay attention to the smell and colour of the coffee, and imagine what it would be like to have a cup of your favourite coffee right at this moment”). In the control condition, participants were presented with a jug of water and asked to imagine being on their favourite holiday (Green et al., 2000). The holiday scenario is an appropriate control match because it is desirable and adaptable to each individual, but is unrelated to either food or beverage.

Manipulation checks and screening questionnaires
After completing the recognition test, participants’ state craving for coffee was assessed retrospectively (as in Kemps et al., 2008). Participants were asked to think back to four time points during the experiment and rate the level of desire for coffee they were experiencing at each: Upon arrival at the laboratory, after the first cue exposure task, after the second cue exposure task, and upon completion of the experiment. Ratings were made on a 100 mm visual analogue scale with anchors of no desire (0%) at the left-hand endpoint and extremely strong desire (100%) at the right-hand endpoint. Scores on a scale of 0–100 were calculated by measuring the distance in mm from the 0% anchor. Retrospective ratings are likely less reliable than real-time ratings, but were used to avoid mentioning the concept of cravings to participants in the control condition before other measures were completed (Kemps et al., 2008).

Measuring the accuracy of metacognitive judgements
We examined the effects of cravings on metacognitive accuracy in several ways. The accuracy of JOLs and FOK judgements was evaluated by examining three distinct properties of metacognitive accuracy: Resolution, calibration, and over/under-confidence. Resolution refers to the extent to which participants can use metacognitive judgements (JOLs or FOK) to distinguish items they will remember from items they will not remember. We measured resolution calculating the Adjusted Normalized Discrimination Index (ANDI; Yaniv et al., 1991). Values range from zero (indicating no discrimination) to one
(perfect discrimination). Note that ANDI is less susceptible to bias than alternative measures of resolution, such as Goodman–Kruskal gamma correlations (e.g., Fleming & Lau, 2014; Masson & Rotello, 2009).

Calibration and over/under-confidence index the degree to which the subjective likelihood of remembering correctly (i.e., JOLS or FOK) corresponds to the objective likelihood of remembering correctly (i.e., accuracy on the test). Calibration curves are constructed by plotting subjective accuracy (JOLS or FOK) against objective accuracy. Perfect calibration occurs when subjective accuracy corresponds to objective accuracy; that is, when items given a JOL or FOK of 20% are actually 20% likely to be remembered correctly, items given a JOL/FOK of 40% are 40% likely to be remembered correctly, items given a JOL/FOK of 90% are actually 90% likely to be remembered correctly, and so forth. The over/under-confidence statistic (O/U) reflects the extent to which average JOLS/FOK judgements are higher or lower than average accuracy, weighted by the number of responses at each level of JOL/FOK (Lichtenstein, Fischhoff, & Phillips, 1982). O/U ranges from −1 to 1, with positive values indicating overconfidence and negative values indicating under-confidence. The calibration statistic (C) is similar to O/U but reflects the absolute weighted difference between mean JOL/FOK and mean accuracy, rather than the directional difference. That is, C indexes the extent to which JOLS/FOK judgements deviate from objective likelihood of remembering correctly, without consideration of whether JOLS/FOK judgements are higher or lower than objective accuracy. The C statistic and ranges from zero (perfect calibration) to one.

**Results**

**Manipulation checks**

The craving group and control group did not differ in average number of coffees consumed per day (craving: $M = 2.0$ v control: $M = 2.0$) or scores on the dependence questionnaire, $t < 1$, Cohen’s $d < 0.06$. The craving group reported a longer average delay since their last coffee ($M = 12$ h, $SD = 3$) than the control group ($M = 5$ h, $SD = 5$), $t (42.45) = 6.37$, $p < .001$, $d = 1.70$. Retrospective state craving ratings indicated that the craving manipulation was successful, with higher ratings for the craving group than the control group across all four time points, as shown in Table 1, $F(1, 53) = 36.62$, $p < .001$.

**Effects of cravings on memory performance**

We tested accuracy on the cued-recall task in two ways, by calculating the proportion of correct responses made out of all 100 trials and the proportion of correct responses made on trials for which an answer was made (i.e., excluding trials on which no answer was made; we term this conditional accuracy). The mean proportion of correct responses across all 100 trials was lower for participants in the craving condition ($M = 0.08$, $SD = 0.13$) than the control condition ($M = 0.20$, $SD = 0.20$), $t (46.2) = 2.52$, $p = .015$, $d = 0.67$. Cravings also reduced conditional accuracy, with a smaller proportion of correct responses in the craving condition ($M = 0.40$, $SD = 0.27$) than control ($M = 0.59$, $SD = 0.26$), $t (53) = 2.64$, $p = .011$, $d = 0.71$.

For the 4AFC recognition task, the mean proportion of correct answers was lower in the craving condition ($M = 0.49$, $SD = 0.19$) than the control condition ($M = 0.63$, $SD = 0.22$), $t (53) = 2.49$, $p = .016$, $d = 0.67$. Thus, caffeine cravings impaired performance on both memory tasks.

**Effects of cravings on metacognition**

Analyses of ANDI, C, and O/U, statistics were conducted with square root transformed scores to correct for skewed distributions, excepting ANDI scores for JOLS and O/U scores for FOK judgements, which were not substantially skewed and not transformed. In cases where no errors or no correct responses were made on a test, we could not compute measures of JOL-ANDI ($n = 3$), FOK-ANDI ($n = 4$), FOK-C ($n = 3$), and FOK-O/U ($n = 3$).

**JOLS**

Table 2 shows means for raw and square root transformed ANDI, C, and O/U statistics for JOLS and FOK judgements in each craving condition. On average, ANDI values did not differ between the craving group and control group, $t < 1$, $p = .529$, $d = 0.18$. Thus, there was no evidence that caffeine cravings affected participants’ ability to use JOLS to distinguish between items they would remember and those they would not.

Cravings did, however, reduce the correspondence between JOLS and objective accuracy (i.e., calibration). The top panel of Figure 1 shows a calibration curve plotting JOLS against accuracy on the cued-recall test. Calibration, indexed by the C statistic, was poorer in the craving condition than the control condition, $t (53) = 2.05$, $p = .046$, $d = 0.55$. However, inspection of the calibration curve suggests that this result may have been partly due to increased overconfidence (O/U) in the craving condition compared to the control condition, despite the latter comparison representing a small and statistically non-significant effect, $t (53) = 1.13$, $p = .263$, $d = 0.31$. Figure 1 (top panel) shows that accuracy was consistently lower for the craving conditions than the control condition at each level of JOL. This pattern, combined with the non-trivial effect size found for the effect of cravings on O/U values,

| Table 1. Descriptive statistics for state coffee craving ratings (scale: 0–100) for the craving and control conditions. |
|-----------------|-----|-----|-----|-----|-----|-----|
|                 | Craving condition |          | Control condition |          |
|                 | $M$  | $SD$| 95% CI | $M$  | $SD$| 95% CI |
| Arrival at lab  | 66  | 29 | [55, 76] | 29  | 25 | [19, 39] |
| Before JOL task | 75  | 27 | [64, 86] | 36  | 29 | [25, 46] |
| Before FOK task | 82  | 19 | [73, 92] | 36  | 28 | [27, 45] |
| End of experiment | 82 | 28 | [71, 92] | 49  | 28 | [39, 60] |
suggests that at least some of the effect of cravings was to increase overconfidence in JOLs. Regardless, although the specific nature of these effects may be open to interpretation, the calibration data indicate that cravings reduced the correspondence between JOLs and cued-recall accuracy.

### Table 2

<table>
<thead>
<tr>
<th></th>
<th>ANDI</th>
<th>C</th>
<th>O/U</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>JOLs</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Craving</td>
<td>.53 (.30)</td>
<td>.06 (.06)</td>
<td>.11 (.14)</td>
</tr>
<tr>
<td>Control</td>
<td>.58 (.26)</td>
<td>.04 (.05)</td>
<td>.07 (.11)</td>
</tr>
<tr>
<td><strong>FOK judgements</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Craving</td>
<td>.06 (.06)</td>
<td>.15 (.11)</td>
<td>–.22 (.23)</td>
</tr>
<tr>
<td>Control</td>
<td>.10 (.06)</td>
<td>.15 (.09)</td>
<td>–.30 (.15)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>√ANDI</th>
<th>√C</th>
<th>√O/U</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>JOLs</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Craving</td>
<td>.68 (.26)</td>
<td>.22 (.13)</td>
<td>.24 (.27)</td>
</tr>
<tr>
<td>Control</td>
<td>.72 (.23)</td>
<td>.15 (.11)</td>
<td>.17 (.24)</td>
</tr>
<tr>
<td><strong>FOK judgements</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Craving</td>
<td>.19 (.16)</td>
<td>.35 (.15)</td>
<td>–.43 (.30)</td>
</tr>
<tr>
<td>Control</td>
<td>.31 (.16)</td>
<td>.38 (.12)</td>
<td>–.54 (.14)</td>
</tr>
</tbody>
</table>

### Discussion

This research extends our understanding of the effects of cravings on cognition in several ways. The effects of cravings on FOK resolution are consistent with the memory constraint hypothesis (Hertzog et al., 2010). That is, cravings impair FOK resolution because they impair strength of memory for target material. By reducing strength of memory for target items, cravings reduce the amount of partial and contextual information that can be accessed when making FOK judgements. This information is crucial to FOK accuracy because it provides valid cues for distinguishing items likely to be recognised from those not likely to be recognised. Thus, by reducing the amount of partial and contextual information accessible during FOKs, cravings reduce resolution of FOK judgements.

Analysis of the FOK resolution and cued-recall memory test scores provided further support for this account. Following Souchay and Isingrini (2012), we examined the correlation between cued-recall memory accuracy and FOK resolution. Consistent with the idea that resolution will be greater at higher levels of memory performance, we found a strong, positive association, r(49) = .63, p < .001. In contrast, cravings did not reduce the resolution of JOLs and, the correlation between JOL resolution and performance on the cued-recall task was trivial, r(50) = −.05, p = .715. Together, these results align with the idea that FOK resolution is dependent on memory strength – because FOK judgements rely on retrieval of partial and contextual information – but JOL resolution is not (Hertzog et al., 2010; Hertzog & Hultsch, 2000; Souchay et al., 2007; Souchay & Isingrini, 2012).

Although cravings did not reduce JOL resolution, they reduced the correspondence between JOLs and cued-recall test scores, as indicated by impaired calibration. As outlined earlier, we suggest that this effect was at least partly due to cravings increasing overconfidence in future cued-recall performance. This would align with many previous demonstrations of the hard-easy
effect, whereby overconfidence tends to be greater under conditions that impair memory performance (Gigerenzer et al., 1991). This interpretation is speculative because the effect of cravings on overconfidence was non-significant. But although the specific nature of these effects might be debated, the calibration data unambiguously indicate that cravings reduced the correspondence between JOLs and cued-recall accuracy.

The effects of cravings on memory performance in the 4AFC task show that cravings can impair performance not only on recall-based tasks but also on recognition-based tasks. This result contrasts previous work which found that cravings impaired recall but had little effect on recognition (Zuj et al., 2015). Why might cravings have impaired recognition in the present study but not in Zuj et al.? The difference in results between these two experiments may be due to the different types of recognition tests used. Zuj et al. used an old–new recognition test that involved a very simple type of decision: Participants were presented with a series of items and asked to determine whether each had appeared previously in the experiment. This type of task can be based on an evaluation of the level of familiarity of each item: Items that are sufficiently familiar can be deemed "old", and those that are not can be deemed "new" (e.g., Yonelinas, 2002). In contrast, the present study used a variation of a 4AFC task. Participants were presented with four response options, comprising one correct option (containing the target word that was paired with the relevant cue word during the study phase) and three lures. This task required participants to not only evaluate the familiarity of response options, but also to evaluate the strength of association between the response options and the cue word (to determine the likelihood that the favoured response options was actually paired with the cue word during the study phase). Thus, the recognition task in the present experiment involved a more complex recognition decisions than the task used by Zuj et al. To the extent that the task used in the present experiment required more cognitive resources than the task used by Zuj et al., there would be more scope for cravings to impair performance. Thus, in light of Zuj et al.’s findings, our results rule out the possibility that cravings do not impair recognition performance, but they are consistent with the notion that cognitively demanding tasks are more susceptible than simple tasks to the effects of cravings (cf. Anderson et al., 1998).

The effects of cravings on cued-recall performance also lead to a new conclusion. Cravings not only reduced accuracy across all trials of the cued-recall test, they also reduced the proportion of accurate responses made on trials for which a response was made (i.e., excluding trials where no response was made). This suggests that cravings impaired participants’ ability to withhold responses that were likely to be incorrect. According to Koriat and Goldsmith’s (1996) model of memory regulation, when attempting to retrieve an answer from memory, people (a) generate a best candidate answer, (b) evaluate the likely accuracy of that answer, then (c) decide whether to report or withhold the answer, based on whether the likely accuracy exceeds a decision criterion. Our results suggest that cravings impair the ability to evaluate the likely accuracy of candidate answers, rather than encouraging a more lenient response criterion for deciding whether to volunteer an answer. Although the latter mechanism could reduce the proportion of correct responses, it would also increase the proportion of trials on which a response was volunteered. This was not the case; in fact, the proportion of trials on which a response was made was lower in the craving condition than the control condition, t(40.78) = 2.85, p = .007, d = 0.76. Thus, it appears that cravings impaired the evaluation of candidate answers generated during the cued-recall task. This unexpected finding provides new evidence that factors that affect memory strength can affect the accuracy of metacognitive judgements.

Our results also point to an interesting conclusion about the effects of cognitive load on memory and metacognition. The effects of cravings on memory performance can be considered a special case of cognitive load: Cravings trigger consumption schemas and the inhibition of these schemas requires cognitive resources, leaving fewer resources available for other cognitive tasks (Kemps et al., 2008; Tiffany, 1990; Tiffany & Conklin, 2000). Previous research shows that some manipulations of cognitive load can reduce memory performance while having minimal impact on the accuracy of metacognitive judgements. For example, Palmer et al. (2013) found that performing a secondary task during the encoding phase reduced memory performance on an eyewitness memory task, but did not reduce the accuracy of post-decision confidence judgements. Palmer et al. reasoned that participants were likely making theory-based adjustments to their metacognitive judgements based on their beliefs about the effects of cognitive load (cf. Koriat & Levy-Sadot, 1999). That is, participants appreciated that their memory was impaired by the secondary task (which involved monitoring a soundtrack and responding to tones of different pitch) and adjusted their metacognitive judgements accordingly. Similar results have been found in studies of multi-tasking (Finley, Benjamin, & McCarley, 2014) and driving behaviour (e.g., Horrey, Lesch, & Garabet, 2008), whereby people appreciate that performance is impaired by a secondary task – such as using a cell phone while driving – and adjust their metacognitive judgements to take this into account (although these adjustments do not always match the degree of actual impairment).

In contrast, our data suggest that people do not appreciate that cravings increase cognitive load and, hence, impair memory performance. In turn, people do not adjust their metacognitive judgements to take this into account. Together, these results suggest that the effects of cognitive load on the accuracy of metacognitive judgements may hinge, in part, on whether people are...
aware of the effects of cognitive load on task performance. If a person is aware that cognitive load is reducing their performance on a task (as with explicit, dual-task manipulations), then they may be able to adjust their metacognitive judgements in an attempt to take this into account. However, under conditions where a person is not aware that cognitive load is affecting their performance (such as covert manipulations of load, including the experience of cravings) there is no apparent basis to make such an adjustment.

**Limitations and conclusions**

It is important to note that we cannot be sure whether the differences observed in memory and metacognition accuracy were due primarily to the operation of cravings at encoding or retrieval. Like previous research (e.g., Kemps et al., 2008), we manipulated cravings via a combination or retrieval. Like previous research (e.g., Kemps & Tiggemann, 2007, 2009, 2015; Kemps, Tiggemann, Woods, & Soekov, 2004; Versland, & Rosenberg, 2007.)

Finally, our focus in this research was on metacognitive judgements about studied material, but the findings may also have implications for other types of metacognitions. For example, some clinical interventions focus on modifying clients’ metacognitions about clinical symptoms (e.g., Wells, 2000), and our results suggest that it would be worthwhile to consider how the effectiveness of these interventions might be influenced by cravings. This is particularly important given that such interventions have been used to treat substance-based and addiction disorders (e.g., Caselli, Gemelli, Spada, & Wells, 2016; Lee, Pohlman, Baker, Ferris, & Kay-Lambkin, 2010).

**Acknowledgements**

We thank Nicole Reid and Frances Parkes for help with data collection.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Funding**

This work was supported by the Australian Research Council under grant DP140103746 to M. Palmer et al.

**References**


Finley, J. R., Benjamin, A. S., & McCarley, J. S. (2014). Metacognition of multitasking: How well do we predict the costs of divided...


