Responses improve the accuracy of confidence judgements in memory tasks.

Marta Siedlecka, Zuzanna Skóra

Jagiellonian University

Borysław Paulewicz

SWPS University of Social Sciences and Humanities

Sonia Fijałkowska

Jagiellonian University

Bert Timmermans

University of Aberdeen

Michał Wierzchoń

Jagiellonian University

Correspondence concerning this article should be addressed to Marta Siedlecka, Institute of Psychology, Jagiellonian University, Kraków, Poland. E-mail: marta.siedlecka@uj.edu.pl

Abstract

How do we assess what we remember? Previous work on metacognition suggests that confidence judgements are more accurate when given after than before a response to a perceptual task. Here we present two experiments that investigate the influence of decision and response on metacognitive accuracy in a memory task so as to establish what kind of information people use to assess their memory content. Participants were asked to remember lists of words and then to decide which of two target words had previously been presented. In both experiments, participants rated their confidence either after or before the response. However, the experiments differed in the amount of information provided for confidence rating. In Experiment 1, before confidence rating, participants were either presented with both target words and asked to decide between them, or they were only presented with a cue (first letter of the subsequent target words). In Experiment 2, participants were always presented with a target word before confidence rating. The results of both experiments showed that although task accuracy correlated with confidence ratings in both conditions, this relationship was weaker when confidence was assessed before response to a memory task. We argue that metacognitive judgements are influenced by processing information that is not available at the time of primary response. We discuss the implications for theories of confidence and metacognition.

Keywords: confidence, decision-making, memory, metacognition, metacognitive awareness

Responses improve the accuracy of confidence judgements in memory tasks.

How do we assess what we perceive and remember? This question taps into the problem of metacognitive access to one's own cognitive states and processes. Multiple theories of metacognition and consciousness claim to provide explanation for how this access occurs (see e.g. Baars, 1997; Cleeremans 2011, Dehaene & Naccache, 2001, Fleming & Daw, 2016; Lau, 2008; Lau & Rosenthal, 2011), however, they are often contradictory and the exact mechanism is still being debated. Here, we focus on the questions concerning what metacognitive judgements about one's knowledge or perceptual experience are based on (an issue that has been vividly discussed in decision models literature, see e.g. Moran, Teodorescu, & Usher, 2015) and whether their mechanisms are the same for different types of content (Baird, Smallwood, Gorgolewski, & Margulies, 2013; McCurdy, Maniscalco, Metcalfe, Liu, de Lange, & Lau, 2013). Specifically, we studied whether response in a memory task improves metacognition, as it has been shown in perceptual (Wierzchoń, Paulewicz, Asanowicz, Timmermans & Cleeremans, 2014) and problem-solving tasks (Siedlecka, Paulewicz, & Wierzchoń, 2016).

Despite the differences between assumed mechanisms, there is an agreement that accessibility can be inferred from metacognitive reports (e.g. Baars, 1997; Dehaene & Naccache, 2001; Lau, 2008; Timmermans, Schilbach, Pasquali, & Cleeremans, 2012). Commonly used quantified forms of those reports are ratings on subjective scales, such as the Perceptual Awareness Scale (PAS), used to measure stimuli visibility (Ramsøy & Overgaard, 2004) and confidence ratings (CR) that assess certainty in a stimulus-related decision (Tunney & Shanks, 2003). The ratings are typically collected during forced-choice tasks (called *type 1* tasks) such as perceptual discrimination or knowledge-based classification (e.g. Dienes & Seth, 2010; Overgaard, Fehl, Mouridsen, Bergholt & Cleeremans, 2008; Sergent &

Dehaene, 2004). Using the ratings (so-called *type 2* measures) one can measure metacognitive accuracy that is the extent to which participants are aware of information they used for *type 1* response. Metacognitive accuracy is therefore measured as a relation between performance level in *type 1* task and subjective ratings (Norman, Price & Jones, 2011; Sandberg Timmermans, Overgaard, & Cleeremans, 2010).

The relation between information that underlies *type 1* response and metacognitive judgement is a subject of debate (e.g. Maniscalco & Lau, 2016; Wierzchoń et al., 2014). One of especially interesting questions concerns the issue of whether metacognitive judgement reflects merely the stimulus-related signal (i.e. its strength) or is based on additional information. Theories of conscious access usually do not explicitly address this question, but the general assumption underlying signal detection-based approaches (e.g. Lau, 2008) or global availability theories (Del Cul, Dehaene, Reyes, Bravo, & Slachevsky, 2009) seems to be that *type 1* responses and metacognitive judgement are based on the same information. This issue has been also investigated in the field of confidence models. A distinction had been made between models assuming single information-collection stage and the ones based on a premise of dual information-collection stage (for a review see e.g. Moran et al., 2015). Single stage models assume that confidence is calculated on the same evidence that led to *type 1* response (Galvin, Podd, Drga, & Whitmore, 2003; Higham, Perfect, & Bruno, 2009; Vickers & Lee, 1998). Proponents of dual stage models challenge this assumption and claim that confidence results from a separate, post-decisional stage of evidence accumulation (e.g. the Two-Stage Dynamic Signal Detection model, Pleskac & Busemeyer, 2010; the Collapsing Confidence Boundary model, Moran et al., 2015). The assumption of an additional stage is supported by the finding that the time taken to rate one's confidence after *type 1* decision predicts reported confidence level. This suggests that confidence is still processed after *type 1* response had been given (Hilgenstock, Weiss, & Witte, 2014; Petrusic

& Baranski, 2003; Yu, Pleskac & Zeigenfuse, 2015). Also, data showing that participants sometimes change their minds after making perceptual decision, even though no more evidence is available (Resulaj, Kiani, Wolpert, & Shadlen, 2009; Van Zandt & Maldonado-Molina, 2004; Yu et al., 2015), support the view that evidence accumulation continues after *type 1* decision and therefore might be used for a subsequent confidence rating. There are also studies showing that choice itself can influence the level of confidence, although the interpretations of this effect differ (Kvam, Pleskac, Yu, & Busemeyer, 2015; Sniezek, Paese, & Switzer III, 1990).

Independently of the issue whether *type 1* decision and metacognitive judgement are based on different amount of evidence, the less frequently asked question concerns whether they use the same source of information. A few recent studies on perception challenge the idea that only stimulus-related information influences the level of confidence. For example, Kiani and colleagues (Kiani, Corthell, & Shalden, 2014) have shown that deliberation time of *type 1* decision negatively correlates with certainty. The relation between confidence and decision time has been found before in number of memory tasks (e.g. Dougherty, Scheck, Nelson, & Narens, 2005; Koriat & Ma'ayan, 2005; Mealor & Dienes, 2013; Petrusic & Baranski, 2003) and was interpreted as reflecting the quality of stimulus or memory trace (e.g. Ratcliff & Starns, 2009; Wixted & Mickes, 2010). However, having controlled stimulus strength, Kiani and colleagues (2014) found that decision time itself inversely influenced subjective confidence, probably as a learnt indicator of task difficulty. Other studies have shown that response-related motor activity and results of error monitoring could also be integrated into perceptual confidence. In the experiment on visual discrimination, participants' certainty in the preceding decision was lowered by transcranial magnetic stimulation (TMS) applied to brain areas associated with motor response opposite to the response chosen by participant (Fleming, Maniscalco, Ko, Amendi, Ro & Lau, 2015). Boldt

& Yeung (2015) have shown that the magnitude of error-related electroencephalography activity that is present after launching an erroneous *type 1* response was negatively related to the level of confidence in the preceding perceptual decision.

To further investigate the issue of sources of information underlying metacognitive judgement, we studied whether *type 1* response could alter metacognitive accuracy (Wierzchoń et al., 2014). In this experiment participants decided whether a backward masked photograph presented a male or female face. Additionally, they gave metacognitive reports using one of four different scales (including CR and PAS), either immediately before or after *type 1* response. Although the accuracy of *type 1* performance did not differ between conditions, the relation between metacognitive judgements and the accuracy of *type 1* response was stronger when the judgement followed discrimination. Therefore, post-decisional metacognitive judgements were more accurate than the pre-decisional ones. A similar effect was found in an anagram-solving task in which participants assessed their confidence either before or after deciding whether a presented word was an anagram solution (Siedlecka et al., 2016). In the light of the previously mentioned experiments, we interpret these results as showing that metacognitive judgements are informed by additional sources of information provided by executing a *type 1* response, such as processing time (Kiani et al., 2014), response-related motor activity (Fleming et al., 2015) or internal accuracy feedback (Boldt & Yeung, 2015).

Although the issue of information fed to the metacognitive judgement seems crucial for understanding the mechanisms of metacognition, there are hardly any studies). This lacuna is quite striking especially as in consciousness research metacognitive judgements are collected before, after or together with *type 1* decision without taking into consideration possible effect of *type 1* response on subjective report (Del Cul, Baillet & Dehaene, 2007; Del Cul et al., 2009; Jachs, Blanco, Grantham-Hill & Soto, 2015; Rounis, Maniscalco,

Rothwell, Passingham, & Lau, 2010; Zehetleitner & Rausch, 2013). Also, in some procedures used to study memory awareness (e.g. *remember-know* procedure; Tulving, 1985), metacognitive judgement preceding *type 1* decision is measured after that decision has been overtly made (Dewhurst, Holmes, Brandt, & Dean, 2006; Mealor & Dienes, 2013). For example, after participants decide whether a presented item was included in the previously studied list, they retrospect whether the decision was based on conscious recollection of an item or on feeling of familiarity towards it (e.g. Dewhurst et al., 2006). Similarly, when retrospective and prospective judgements are measured in memory studies, the latter are reported after retrieval attempts (Dougherty et al., 2005; Schnyer, Verfaellie, Alexander, LaFleche, Nicholls, & Kaszniak, 2004).

In order to propose a general mechanism underlying metacognitive judgements one needs to integrate data from studies using different types of content. There have been a few attempts to compare the accuracy and neural correlates of metacognitive judgements in perceptual and memory domains, but the results are not clear-cut. For example, although in some experiments metacognitive accuracy did not differ between perceptual and memory tasks (McCurdy et al., 2013, but see: Baird et al., 2013), it seemed to have distinct neural substrates in each domain (Baird et al., 2013; McCurdy et al., 2013). Following the quest for understanding universal mechanisms of metacognitive awareness we present a set of experiments aiming to test the influence of a *type 1* response on metacognitive accuracy in tasks that involve memory. We designed experiments that were procedurally similar to the perceptual task used by Wierzchoń and colleagues (2014) in that participants rated their metacognition either before or after a two-alternative forced choice (2-AFC) *type 1* task.

We conducted two experiments to compare metacognitive accuracy between conditions in which confidence ratings directly followed or preceded a *type 1* response. Each experiment contained the two conditions, but they differed in respect to *type 1* task. In

Experiment 1, participants were asked to decide which of the two target words had been presented previously. In Experiment 2, participants responded whether a target word had been presented before. The experimental procedures allowed for varying amounts of decision-related information to be available to participants before they reported their confidence. In Experiment 1, we investigated whether confidence ratings given after a decision were more accurate than when given before a decision. In this experiment, participants either responded first to the target words ("which of the two words has been presented before?") and then rated confidence in that response, or, in the second condition, they were asked to rate their confidence in a subsequent response before even seeing the target words (they were presented only with a cue: the first letter of the word to recall). This procedure allowed us to create conditions that differed in the presence of *type 1* decisions, as it prevented participants in the second condition from deciding between the two target words prior to confidence rating. However, as a result, the two conditions differed also in respect to the memory process (recognition and recall) preceding the metacognitive judgement. Therefore, in Experiment 2 the target word was always presented at the beginning of the trial and conditions differed only in the presence of a *type 1* (motor) response before confidence rating.

We expected that metacognitive accuracy would be higher for confidence ratings given after a *type 1* decision and after a *type 1* response; this would support the view that carrying out a *type 1* response provides additional information to that already available at the time of that response. This additional information processing might also be inferred from temporal characteristics of response and confidence rating. As metacognitive judgement is at least partially based on the same evidence as *type 1* responses (i.e. memory trace strength), we expected that RT in a memory task would be negatively correlated with the level of confidence (Dougherty et al., 2005; Mealor & Dienes, 2013; Petrusic & Baranski, 2003;

Pleskac & Busemeyer, 2010). This relation, however, should be stronger in the case of ratings given after *type 1* responses, as they could be informed by the amount of time used to respond. Additionally, we hypothesized that the time it takes to give confidence rating would be negatively correlated with the level of confidence (Hilgenstock et al., 2014; Petrusic & Baranski, 2003).

We also investigated the relation between the timing of metacognitive judgement and metacognitive accuracy. There are two ways in which the duration of assessing confidence could relate to metacognitive accuracy. On one hand, some researchers suggest that task difficulty affects metacognition (Galvin et al., 2003); therefore, if rating time reflects the difficulty of forming a metacognitive judgement, the later the metacognitive rating, the lower the metacognitive accuracy should be in both conditions. Alternatively, we could expect a positive correlation between metacognitive accuracy and post-decisional confidence judgement time, which would reflect the time needed to integrate additional information from different sources, such as decision time or results of error monitoring (e.g. Yu et al., 2015).

Experiment 1

The aim of Experiment 1 was to assess whether the order of *type 1* decision and confidence rating influences metacognitive accuracy.

Methods

Participants. Thirty-eight volunteers (27 women) aged 18-30 ($M = 22.47$, $SD = 2.63$) took part in the experiment in return for a small payment (about 10 PLN, an equivalent of 2.5 EUR). All participants had normal or corrected-to-normal vision and gave written consent to participation in the study. The protocol for all experiments was approved by an institutional ethical board and was carried out in accordance with the provisions of the World Medical Association Declaration of Helsinki.

Materials. The experiment was run on PC computers using E-Prime (version 2.0). The stimuli were 240 Polish six-to-eight letter words (maximum three syllables) chosen from a subtitle-based frequency list (Mandera, Keuleers, Wodniecka, & Brysbaert, 2014). Because the participants' task was to decide which of the two words had appeared on the list (McCurdy et al., 2013), two alternative versions of the word lists were created. Therefore, each word could be an "old" or "new" word, depending on the list presented to a given participant. Each of the six lists contained 20 words, each starting with a different letter. The confidence scale was associated with either the question "How confident are you that you will make the right decision?", or "How confident are you that you made the right decision?" The alternatives were "I am guessing", "I am not confident", "I am quite confident", and "I am very confident" (Dienes & Perner, 2004).

Procedure. The experiment took place in a computer laboratory. Participants were randomly assigned to one of two conditions that differed in respect to the task order: either the response to the target words (*type 1*) preceded the confidence rating (R-CR), or the confidence rating preceded the word-related response (CR-R). There were 6 experimental blocks, each started with a list of words presented at the centre of the screen for 20 s and was followed by a distracting task consisting of 6 algebra equations, to which participants responded by typing the outcome number. Twenty experimental trials followed, each starting with a cue presented in the centre of the screen for 500 ms, being the first letter of a two subsequently presented words. Then participants either performed the *type 1* task or rated their confidence. During the *type 1* task two words, starting with the same letter, were presented on the left and right side of the screen. Participants were asked to decide which word was presented on the previously studied list using keys "1" and "2" on the numeric keyboard with their right hand. In the confidence rating task, participants were asked to assess their confidence in choosing the right word. Participants used keys "1", "2", "3", "4"

on the keyboard with their left hand. The time available for both, response in *type 1* task and confidence rating was 5 s. After each trial a blank screen was presented for 1 s. At the beginning of the procedure participants were presented with one training list with 5 words and 5 subsequent recognition trials. The procedure is presented on Figure 1 (part A).

Data analysis

Metacognitive accuracy was operationalized as the relationship between *type 1* performance and confidence rating (Norman et al., 2011; Sandberg et al., 2010), and precisely as the extent to which confidence ratings predict *type 1* response accuracy. This relation was analysed using logistic regression, which is the correct model for predicting binary outcomes (Norman & Price, 2015). Logistic regression has important advantages over popular approaches based on signal detection theory: (1) the mixed model framework allows us to answer several statistical questions and to control for the random effect of subjects and stimuli in the context of a single comprehensive analysis; (2) mixed models tolerate unbalanced designs and relatively small trial numbers; (3) logistic regression is not based on the assumption that the information available for the metacognitive judgement is a function only of the information available for the *type 1* decision.

The mixed logistic regression models were fitted using the lme4 package in the R Statistical Environment (Bates, Maechler, Bolker, & Walker, 2015; R Core Team, 2015) using standard (0/1) contrast coding. In Experiment 1, our main model included the following fixed effects: confidence rating (4 levels), order (2 levels), their interactions, as well as the random effects of participant-specific intercept, word, and slope. CRs were centred on the lowest value ("guessing") and the basic condition was the post-decisional CR (R-CR). Therefore, the regression slope reflects the relation between *type 1* task accuracy and confidence rating (metacognitive accuracy) while the intercept informs about performance

level when participants report guessing. Statistical significance was assessed with the Wald test. The data were not analysed prior to completion of the collecting stage.

Results

Two participants' data were excluded from analysis, one due to poor performance (34% accuracy) and one whose variance of scale usage (the number of scale points used) was three standard deviations below average. Next, we removed 44 omitted trials. We found no differences between conditions neither in the *type 1* task accuracy ($z = -0.94, p = .35$) nor with respect to the response bias ($z = 1.67, p = .09$). The general descriptive statistics are presented in Table 1 and Table 2. The frequency with which participants used each scale point and the average accuracy for each scale point are presented on Figure 2.

To analyse the effect of decision-confidence rating order on metacognitive accuracy (that is the relation between *type 1* accuracy and confidence rating) we fitted a mixed logistic regression model defined above. The results are presented in Table 3. The first line (intercept) represents *type 1* response accuracy in the baseline condition (R-CR, lowest rating = "guessing") on the logit scale; so the nonsignificant intercept indicates that when participants claimed they were guessing, their response accuracy did not differ significantly from chance level. The second coefficient estimates the difference between conditions (R-CR and CR-R) in average accuracy for the lowest scale point. This difference is statistically significant and shows that, when reporting guessing, participants in CR-R performed the memory task with higher accuracy than in R-CR. The third line estimates the statistically significant relationship between confidence and accuracy (in R-CR, this relationship is statistically significant). The last coefficient estimates the effect of condition on the confidence rating slopes. The relationship between confidence and accuracy was weaker in the CR-R condition than in the R-CR condition. Additionally, we analysed the relation between confidence and accuracy within CR-R, which was also statistically significant ($z = 2.23, p = .02$).

The observed advantage of R-CR in metacognitive accuracy over CR-R could be attributed to the difference in the amount of time available for making metacognitive judgements (the time between the presentation of a letter cue and confidence rating). To investigate this possibility we created a new variable, CR time, which in the case of R-CR was equal to the sum of *type 1* response RT and confidence rating RT, and in CR-R was equal to confidence rating RT. We did not find statistically significant relationship between CR time and metacognitive accuracy in either condition (R-CR: $z = -1.81$, $p = .07$, CR-R: $z = 0.7$, $p = .49$).

Lastly, to explore between-condition differences in relations between *type* 1 accuracy, confidence level and temporal characteristics of both, we performed a series of analyses using mixed logistic regression models (dependent variable: accuracy) and linear mixed models (dependent variable: confidence rating). Participants chose the correct word quicker in both conditions (R-CR: $z = -12.29$, $p < .001$; CR-R: $z = -9.73$, $p < .001$), and we found no differences in the strength of this relationship between conditions ($z = 0.87$, $p = .38$). Shorter *type 1* response RT was associated with higher confidence in R-CR ($z = -26.68$, $p < .001$), but not CR-R ($z = -1.79$, $p = .08$). The difference between the conditions was statistically significant ($z = 4.13$, $p < .001$). Similarly, when participants rated their confidence after *type 1* response, their metacognition was more accurate when the *type 1* response RT was shorter (R-CR: $z = -4.45$, $p < .001$). When confidence was rated before the response, we found no relation between RT and metacognitive accuracy (CR-R: $z = -.66$, $p = .51$). Additionally, the confidence RT correlated with the reported confidence level in both conditions (the longer confidence RT the lower confidence, R-CR: $z = -11.72$, $p < .001$, CR-R: $z = -2.62$, $p = .01$) although the relationship was significantly stronger in R-CR ($z = 8.23$, $p < .001$). However, we did not find a relation between confidence RT and metacognitive accuracy in either condition (R-CR: $z = .55$, $p = .58$, CR-R: $z = .86$, $p = .39$).

Discussion

The results of Experiment 1 replicated the main finding from the visual awareness study (Wierzchoń et al., 2014) showing that metacognitive judgements are more accurate when measured after *type 1* response. The relation between confidence level and *type 1* accuracy was significantly stronger in R-CR condition. The effect could not be explained neither by the difference in difficulty between the two conditions nor by the amount of time available for metacognitive judgement. The comparison of average accuracy for the lowest scale point showed that participants in CR-R, when claiming that they would be guessing, actually performed better than in the R-CR condition (in which the accuracy did not differ significantly from chance level). The requirement of assessing one's confidence in future decision probably yields to a more cautious strategy than in situation when confidence about past decision is assessed. Although in CR-R the "guessing" rating was not chosen more often that other scale points we can see that participants used middle scale points more often than the highest and the lowest end of the scale. On the contrary participants in R-CR used the highest scale points most often.

The analysis of the relation between *type 1* RT, confidence rating RT and confidence level replicated some previous findings but also showed further differences between conditions. The relation between *type 1* RT and the level of confidence as well as metacognitive accuracy suggests that information used for *type 1* decision influences also confidence rating (Dougherty et al., 2005; Pleskac & Busemeyer, 2010) as *type 1* RT is usually interpreted as reflecting the quality of evidence (e.g. memory strength, Ratcliff & Starns, 2009; Wixted & Mickes, 2010). However, as expected, this relation was significantly stronger in R-CR condition. Therefore the response time in *type 1* task could inform confidence rating - for example, via a heuristic that response time reflects task difficulty (Kiani et al., 2014).

The negative relation between the latency of confidence rating and the level of confidence has been found previously in perceptual and memory studies (e.g. Dewhurst et al., 2006; Mealor & Dienes, 2013; Baranski & Petrusic, 1998; Petrusic & Baranski, 2003; Pleskac & Busemeyer, 2010; Yu et al., 2015), and suggests that confidence is processed, at least partially, separately from *type 1* response (e.g. Pleskac & Busemeyer, 2010). Conditions differed in terms of those relationships: it was stronger in R-CR. Those results probably stem from the fact that in CR-R condition CR RT and *type 1* RT are partly overlapping (CR RT includes the time of attempting to access the word).

Experiment 2

The aim of Experiment 2 was to assess whether the order of *type 1* response and confidence rating influences metacognitive accuracy. The results of Experiment 1 supports the hypothesis that *type 1* decision can influence metacognitive accuracy. However, one difficulty with interpreting the results is that participants in the CR-R condition were not presented with the two target words before rating their confidence. This procedure was necessary to minimize the chance of participants making an actual decision before confidence rating in CR-R condition. Although in Experiment 1 we found no difference in *type 1* accuracy or bias (proportion of "left" or "right" responses) between the conditions, it is possible that the discrepancies in metacognitive accuracy stemmed from the difference between processes leading to *type 1* response. For example, participants recognized a word in the R-CR condition but had to recall it before reporting confidence in CR-R. While post-decisional ratings could be based on the outcome of word recognition, the pre-decisional metacognition could have been informed by a recall attempt following the letter cue.

In Experiment 2 participants were always presented with a target word before being asked for a *type 1* response or confidence rating, and therefore performed the same recognition task in both conditions. This procedure helped us to avoid aforementioned

problems, but allowed participants to make their world-related decision before rating confidence also in the condition in which the rating preceded the *type 1* response.

Method

  Participants. Fifty-five volunteers (8 males) aged 19-30 ($M = 20.31$, $SD = 1.64$), took part in the experiment in return for a small payment (about 10 PLN, an equivalent of 2.5 EUR). All participants had normal or corrected-to-normal vision and gave written consent to participation in the study.

  Materials. Materials were the same as in Experiment 1.

  Procedure. The general procedure was similar to the one used in Experiment 1. Participants were randomly assigned to one of two conditions that differed with respect to the *type* 1 response and confidence rating order. Contrary to the previous experiment, here the target word was always presented before word-related response and confidence rating were required. Participants were presented with just one target word and then asked to answer whether it had been previously presented on the list. We changed this detail so that participants did not have to keep two words and their locations (left or right) in memory until the moment of response. The letter cue was replaced with a fixation point and the target word was presented for 500 ms. Also, the time available for responses was decreased to 3 seconds. The procedure is presented in Figure 1 (part B).

Results

  None of the participants were excluded from analysis due to poor performance or low variance of scale usage. We removed 107 omitted trials. We did not find significant differences between conditions in *type 1* task accuracy ($z = -1.09$, $p = .29$), nor with respect to response bias ($z = -0.56$, $p = .57$). The general descriptive statistics are presented in Table 4 and Table 5. The frequency with which participants used each scale point and the average accuracy for each scale point are presented on Figure 3.

Similarly to the model used in Experiment 1, the first line in Table 6 (intercept) describes the difference (on the logit scale) between accuracy and chance level in the baseline condition (Order R-CR, lowest rating = "guessing"). The second coefficient estimates the difference between conditions (R-CR and CR-R) in average accuracy for the lowest rating ("guessing"). We did not find such a difference between conditions. The third line estimates the relationship between confidence rating and accuracy (in the baseline condition, R-CR). This relationship is statistically significant. The last coefficient estimates the effect of condition on the confidence rating slopes. Confidence ratings were less strongly related to accuracy when given before a *type 1* response (the test was two-tailed, therefore for the directional hypothesis that we were testing, $p = .029$). Additionally, we calculated the correlation between confidence ratings and accuracy within the CR-R condition; it was statistically significant ($z = 10.19, p < .001$).

Since the observed advantage of R-CR over CR-R in metacognitive accuracy could be attributed to the difference in the amount of time available for giving a confidence rating, we created a variable called CR time: in condition CR-R it equalled confidence rating RT and in condition R-CR it equalled *type 1* RT plus confidence rating RT. We found a statistically significant relation between CR time and metacognitive accuracy (R-CR: $z = -2.83, p = .005$. CR-R: $z = -2.43, p = .02$): the later the confidence rating was given, the lower the metacognitive accuracy was. We did not find a significant difference in the strength of this relationship between the two conditions ($z = 0.13, p = 0.9$).

To explore the differences in the relation between *type 1* response accuracy, confidence rating, and the temporal characteristics of both, we performed a series of analyses using mixed logistic regression models (dependent variable: accuracy) and linear mixed models (dependent variable: confidence rating). Participants chose the correct word more quickly in both conditions (R-CR: $z = -7.04 , p < .001$; CR-R: $z = -5.38, p < .001$) and we

found no differences between conditions ($z = 0.4$, $p = .69$). Shorter *type 1* response RT was associated with higher confidence in R-CR ($z = -8.81$, $p < .001$) and in CR-R ($z = -10.31$, $p < .001$). We did not observe a significant difference between the conditions in the strength of this relationship ($z = -1.54$, $p < .13$). Similarly, the quicker the *type 1* response, the more accurate the metacognition in both conditions (R-CR: $z = -3.33$, $p < .001$, CR-R: $z = -3.12$, $p = .001$). We found no significant difference between the conditions ($z = -0.55$, $p = 0.58$). Lastly, the longer the confidence RT, the lower the confidence level in both conditions (R-CR: $z = -6.69$, $p < .001$, CR-R: $z = -6.1$, $p < .001$, no significant difference between conditions: $z = 0.88$, $p < .39$). However we found no significant relationship between confidence RT and metacognitive accuracy in R-CR ($z = -1.1$, $p = .27$), while in the CR-R condition the confidence ratings that were given later were associated with lower metacognitive accuracy ($z = -2.37$, $p = .02$).

Discussion

The results of Experiment 2 replicated the main finding from Experiment 1, showing that metacognitive accuracy is higher when metacognitive judgement follows *type 1* response. The effect was smaller than in the previous experiment, suggesting, that in Experiment 1 the observed difference in metacognitive accuracy could have been enlarged by the differences in the amount of decision-related information available before confidence rating (recognition in R-CR vs. recall in CR-R condition). The results of Experiment 2 suggest that the presence of *type 1* response before metacognitive judgement increases metacognitive accuracy, even though a *type 1* decision could have been made before confidence rating. Interestingly, the time of confidence judgement (measured from the beginning of a trial) was related to the metacognitive accuracy, but in a way that could not explain our main effect. Specifically, the later the confidence rating, the lower the metacognitive accuracy in both conditions.

Contrary to Experiment 1, we found no differences between conditions in *type 1* accuracy level for the lowest scale rating ("guessing"). Also, the frequency of the chosen scale points was similar in both conditions. This supports our interpretation of the results of the previous experiment. When participants were asked to rate their confidence in their future decision when there was limited data (they had not seen the target words yet), they engaged a more careful strategy compared to the condition in which they knew the target words before assessing the confidence of their subsequent decision.

The analysis of the relation between *type 1* response RT, confidence rating RT and confidence level showed further differences between the experiments. Contrary to Experiment 1, we found a relation between *type 1* response RT and the level of confidence and a relation between *type 1* response RT and metacognitive accuracy in both conditions. Therefore, presenting target words at the beginning of trials might have diminished the differences between conditions. After seeing the word participants might have attempted to make their *type 1* decisions and probably got some sense of trial difficulty, experiencing fluency or item familiarity (Metcalfe, Schwartz, & Joaquin, 1993; Koriat & Ma'ayan, 2005). It has been shown that pre-judgement retrieval fluency (that is, its latency and success) correlates with metacognitive judgements about past and future memory performance (Dougherty et al., 2005; Matey, Dunlosky, & Guttentag, 2001; Kelley & Lindsay, 1993).

Similarly to Experiment 1, confidence rating RT was negatively related to the confidence level in both conditions, which again supports the hypothesis that confidence is processed, at least partially, separately from *type 1* responses (e.g. Pleskac & Busemeyer, 2010; Yu et al., 2015). Also, similarly to the Experiment 1, we did not observe a significant relationship between confidence rating RTs and metacognitive accuracy in the R-CR condition, although we found a negative relationship in CR-R. Therefore we are not able to determine whether the time taken to judge post-decisional confidence is associated with the

difficulty of processing confidence (decreased metacognitive accuracy for longer confidence RTs) or with integrating additional information (increased metacognitive accuracy for longer confidence RTs). This is interesting in the context of studies showing that increasing the time between decision and confidence rating improves confidence resolution (the degree to which confidence judgements discriminate between correct and incorrect choices, Yu et al., 2015). In the study of Yu and colleagues (2015), participants were prompted to rate their confidence at given moments; when the post-decisional time increased, the confidence in incorrect decisions decreased. It seems that when a state of confidence in each time interval is measured it might reflect additional information integrated from the error-monitoring system. On the other hand, when participants are free to choose the moment of rating (within a time limit, as in our experiments), the process becomes more noisy and contaminated with other processes (e.g. an evidence searching strategy), therefore it becomes difficult to determine its relation to the accuracy.

General discussion

The main finding of the presented experiments is that metacognitive assessments are more accurate when given after rather than before *type 1* decision. Metacognitive accuracy does not improve with time therefore the effect could not be explained away only by the difference in interval between letter cue and confidence rating. The data also showed that the level of confidence correlates not only with *type 1* response time but also with the time of the rating itself. The findings are consistent across the two experiments, even though the procedures differed in a several ways. In Experiment 1 participants performed different *type 1* tasks in each condition (recognition in R-CR and recall in CR-R) while in Experiment 2 the same recognition task was used in both conditions. Also, in Experiment 1 participants saw a target and could make their *type 1* decision before confidence judgement only in the R-CR condition; in Experiment 2 the target was always presented before *type 1* response and

confidence rating, therefore conditions differed only in the order of these two responses. Together with data on the relation between the level of confidence and temporal characteristics of *type 1* response and confidence rating, our results suggest that metacognitive judgements are based on information integrated from multiple sources; this could include external cues, i.e. those available in the task itself, or internal cues such as memory trace strength, decision and response given by a participant. The outcome of the experiments also support the view that processes underlying subjective ratings are universal, as the influence of response on metacognitive accuracy was also found in a perceptual study (Wierzchoń at al, 2014) and problem-solving task (Siedlecka et al., 2016).

The results interpreted in the context of confidence models support the hypothesis that metacognitive reports are based on different amount of information than available at the time of *type 1* response (Graziano, Parra, & Sigman, 2015; Hilgenstock et al., 2014; Petrusic & Baranski, 2003; Pleskac & Busemeyer, 2010). Although those models apply to post-decisional confidence, the results showing the difference between post and pre-decisional and post and pre-response confidence ratings stand against single information-collection stage theories (Vickers & Lee, 1998; Higham et al., 2009). Assuming that confidence is set at the time of a *type 1* decision, one could expect larger latency of confidence rating in the CR-R condition because of the inclusion of both the attempt to recall/recognize the word and the confidence processing time; however, the correlation between confidence and *type 1* accuracy should be equally strong in both conditions. This was not the case in our study. Moreover, we found negative correlations between confidence rating RT and the level of confidence, usually interpreted as indicating a separate stage of evidence accumulation (Hilgenstock et al., 2014; Petrusic & Baranski, 2003).

The data from our experiments also suggest that metacognitive judgements are based on information integrated from different sources. Although, our study was not designed to

directly identify those sources, the data indicates that they include *type 1* decision and response characteristics. The analysis of reaction times has shown that delayed *type 1* responses were associated with a lower level of metacognitive awareness (in Experiment 1 only for post-decisional confidence rating). Since response time is thought to be an index of task difficulty (e.g. Wixted & Mickes, 2010), this result is in line with data showing that the ease or success of the word retrieval attempt (Dougherty et al., 2005; Kelley & Lindsay, 1993; Koriat & Ma'ayan, 2005) and the duration of actual decision (Kiani et al., 2014) informs metacognitive judgement. Similarly, the aforementioned studies on response-related activity of the premotor cortex (Fleming et al., 2015), results of response-related error monitoring (Boldt & Yeung, 2015; Scheffers & Coles, 2000) and interceptive feedback (Wessel, Danielmeier, & Ullsperger, 2011) point to the interpretation that response-related characteristics are a crucial source of information for metacognitive assessment. Interestingly, Fleming and Daw (2017) have recently proposed a second-order model of metacognition in which one's own actions contribute to confidence reports (Fleming & Daw, 2017). The predictions of the model were compared to data from our experiment on an anagram task (Siedlecka et al., 2016) which displayed similar patterns.

Considering alternative explanations for the influence of *type* 1 decision and response on metacognitive accuracy we should mention studies showing the effect of choice on confidence calibration (understood as the probability of being correct that people assign to their answers, relative to the actual proportion correct, Kvam et al., 2015; Sniezek et al., 1990). For example, the quantum confidence model assumes that when a decision is made the indefinite state of evidence changes into a definite choice state, reducing a person's uncertainty (Kvam, et al., 2015). The support for this model comes from experiments showing that participants are less overconfident in the condition in which *type 1* response precedes confidence rating compared to the condition in which participants simultaneously

reported both *type 1* decision and confidence (Kvam et al., 2015). On the other hand, Śniezek and colleagues (1990) suggested that explicit choice reduces overconfidence due to more cognitive processing related to decision-making. The data from our experiments are not in line with the results from the aforementioned studies: in Experiment 1 we actually observed underconfidence (above-chance performance level for "guessing") in the condition in which *type 1* decision followed confidence. This difference between the results might be yet another example of the importance of the sequential relation between *type 1* decision, response and confidence judgement. For example, when participants are asked to report decision and confidence on the same scale (i.e. "certain that A - certain that B", Kvam et al. 2015), although confidence does not follow *type 1* response, it is probably processed differently to a situation in which no decision was made prior to or simultaneously with confidence rating.

In future studies we should try to separate the decision itself and its motor execution in a single experiment. Although studies on perceptual confidence point to the role of neural activity related to motor response (Boldt & Yeung, 2015; Fleming et al., 2015), we cannot exclude the possibility that covert decisions also affect metacognitive accuracy. However, it is hard to address this issue experimentally. When participants are asked to make a decision without a response then to measure the relation between *type 1* accuracy and metacognitive judgement we have to trust them to remember and report the same decision afterwards (after metacognitive judgement), even if they change their minds in the meantime. At the same time, it is difficult to design a condition in which participants carry out motor response to a task without making a decision first. One way to achieve this is to induce a task irrelevant motor reaction that might overlap with stimuli-related responses (Siedlecka, Hobot, Skóra, Paulewicz, Timmermans, & Wierzchoń, in preparation).

To conclude, our results indicate that response in a memory task improves metacognitive assessment of related memory content. However, further research is needed to

determine the mechanisms of metacognitive assessments and the factors influencing its accuracy. Also, the observed discrepancy between metacognitive accuracy for confidence rating given after and before *type 1* response adds to the empirical hurdles that confidence models should overcome (Pleskac & Busemeyer, 2011; Moran et al., 2015). Apart from the theoretical implications, there are also some practical outcomes. We suggest that subjective reports proceeding and following *type 1* responses should not be used interchangeably.

References

Baars, B. J. (1997). *In the theatre of consciousness: The workspace of the mind*. Oxford: Oxford University Press.

Baird, B., Smallwood, J., Gorgolewski, K. J., & Margulies, D. S. (2013). Medial and lateral networks in anterior prefrontal cortex support metacognitive ability for memory and perception. *The Journal of Neuroscience*, *33*(42), 16657-16665.

Baranski, J. V., & Petrusic, W. M. (1998). Probing the locus of confidence judgements: experiments on the time to determine confidence. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 929-945.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1-48.

Boldt, A., & Yeung, N. (2015). Shared Neural Markers of Decision Confidence and Error Detection. *The Journal of Neuroscience*, *35*(8), 3478-3484.

Cleeremans, A. (2011). The radical plasticity thesis: how the brain learns to be conscious. *Frontiers in Psychology, 2*(36), 86.

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition, 79*(1), 1-37.

Del Cul, A., Baillet, S., & Dehaene, S. (2007). Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS Biology*, *5*(10), e260.

Del Cul, A., Dehaene, S., Reyes, P., Bravo, E., & Slachevsky, A. (2009). Causal role of prefrontal cortex in the threshold for access to consciousness. *Brain*, *132*(9), 2531-2540.

Dewhurst, S. A., Holmes, S. J., Brandt, K. R., & Dean, G. M. (2006). Measuring the

   speed of the conscious components of recognition memory: Remembering is

   faster than knowing. *Consciousness and Cognition*, *15*(1), 147-162.

Dienes, Z., & Perner, J. (2004). Assumptions of a subjective measure of

   consciousness: Three mappings. In R. Gennaro (Ed.), *Higher order theories of*

   consciousness (pp 173-199). Amsterdam: John Benjamins Publishers.

Dienes, Z., & Seth, A. (2010). Gambling on the unconscious: A comparison of

   wagering and confidence ratings as measures of awareness in an artificial

   grammar task. *Consciousness and Cognition, 19*(2), 674–681.

Dougherty, M. R., Scheck, P., Nelson, T. O., & Narens, L. (2005). Using the past to

   predict the future. *Memory & Cognition*, *33*(6), 1096-1115.

Fleming, S. M., & Daw, N. D. (2017). Self-evaluation of decision-making: A general

   Bayesian framework for metacognitive computation. *Psychological Review,*

   124(1), 91-114.

Fleming, S. M., Maniscalco, B., Ko, Y., Amendi, N., Ro, T., & Lau, H. (2015).

   Action-specific disruption of perceptual confidence. *Psychological Science*,

   26(1), 89-98.

Galvin, S. J., Podd, J. V., Drga, V., & Whitmore, J. (2003). Type 2 tasks in the theory

   of signal detectability: Discrimination between correct and incorrect decisions.

   Psychonomic Bulletin & Review, 10(4), 843-876.

Graziano, M., Parra, L. C., & Sigman, M. (2015). Neural Correlates of Perceived

   Confidence in a Partial Report Paradigm, *Journal of Cognitive Neuroscience*,

   27(6), 1090-1103.

Higham, P. A., Perfect, T. J. & Bruno, D. (2009). Investigating strength and frequency effects in recognition memory using type-2 signal detection theory. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *35*, 57-80.

Hilgenstock, R., Weiss, T., & Witte, O. W. (2014). You'd Better Think Twice: Post-Decision Perceptual Confidence. *NeuroImage*, *99*, 323-331.

Jachs, B., Blanco, M. J., Grantham-Hill, S., & Soto, D. (2015). On the independence of visual awareness and metacognition: A signal detection theoretic analysis. *Journal of Experimental Psychology: Human Perception and Performance 41*(2), 269-276.

Kelley, C. M., & Lindsay, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language*, *32*(1), 1-24.

Kiani, R., Corthell, L., & Shadlen, M. N. (2014). Choice certainty is informed by both evidence and decision time. *Neuron*, *84*(6), 1329-1342.

Koriat, A., & Ma'ayan, H. (2005). The effects of encoding fluency and retrieval fluency on judgements of learning. *Journal of Memory and Language*, *52*(4), 478-492.

Kvam, P. D., Pleskac, T. J., Yu, S., & Busemeyer, J. R. (2015). Interference effects of choice on confidence: Quantum characteristics of evidence accumulation. *Proceedings of the National Academy of Sciences, 112(34)*, 10645-10650.

Lau, H. C. (2008). A higher order Bayesian decision theory of consciousness. *Progress in Brain Research, 168*, 35-48.

Lau, H., & Rosenthal, D. (2011). Empirical support for higher-order theories of conscious awareness. *Trends in Cognitive Sciences*, *15*(8), 365-373.

Mandera, P., Keuleers, E., Wodniecka, Z., & Brysbaert, M. (2014). Subtlex-pl: subtitle-based word frequency estimates for Polish. *Behavior Research Methods*, *47*(2), 1-13.

Maniscalco, B., & Lau, H. (2016). The signal processing architecture underlying subjective reports of sensory awareness. *Neuroscience of Consciousness*, *2016*(1), niw002.

Matvey, G., Dunlosky, J., and Guttentag, R. (2001). Fluency of retrieval at study affects judgements of learning (JOLs): an analytic or nonanalytical basis for JOLs. *Memory & Cognition, 29*, 222-233.

McCurdy, L. Y., Maniscalco, B., Metcalfe, J., Liu, K. Y., de Lange, F. P., & Lau, H. (2013). Anatomical coupling between distinct metacognitive systems for memory and visual perception. *The Journal of Neuroscience, 33*(5), 1897-1906.

Mealor, A. D., & Dienes, Z. (2013). The speed of metacognition: Taking time to get to know one's structural knowledge. *Consciousness and Cognition*, *22*(1), 123-136.

Metcalfe, J., Schwartz, B. L., & Joaquim, S. G. (1993). The cue-familiarity heuristic in metacognition. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 19*, 851-561.

Moran, R., Teodorescu, A. R., & Usher, M. (2015). Post choice information integration as a causal determinant of confidence: Novel data and a computational account. *Cognitive Psychology, 78*, 99-147.

Norman, E., & Price, M. C. (2015). Measuring consciousness with confidence ratings. In M. Overgaard (Ed.) *Behavioural Methods in Consciousness Research* (pp. 159-180). Oxford: Oxford University Press.

Norman, E., Price, M. C., & Jones, E. (2011). Measuring strategic control in artificial grammar learning. *Consciousness and Cognition*, *20*(4), 1920-1929.

Overgaard, M., Fehl, K., Mouridsen, K., Bergholt, B., & Cleeremans, A. (2008). Seeing without seeing? Degraded conscious vision in a blindsight patient. *PLoS One, 3*(8), e3028.

Petrusic, W. M., & Baranski, J. V. (2003). Judging confidence influences decision processing in comparative judgements. *Psychonomic Bulletin & Review*, *10*(1), 177-183.

Pleskac, T. J., & Busemeyer, J. R. (2010). Two-stage dynamic signal detection: a theory of choice, decision time, and confidence. *Psychological Review, 117*, 864-901.

R Core Team (2015). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna.

Ramsøy, T. Z., & Overgaard, M. (2004). Introspection and subliminal perception. *Phenomenology and the Cognitive Sciences, 3*(1), 1–23.

Ratcliff, R., & Starns, J. J. (2009). Modeling confidence and response time in recognition memory. *Psychological Review*, *116*(1), 59-83.

Resulaj, A., Kiani, R., Wolpert, D. M., & Shadlen, M. N. (2009). Changes of mind in decision-making. *Nature, 461*(7261), 263-266.

Rounis, E., Maniscalco, B., Rothwell, J. C., Passingham, R. E., & Lau, H. (2010). Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs metacognitive visual awareness. *Cognitive Neuroscience, 1(3)*, 165–175.

Sandberg, K., Timmermans, B., Overgaard, M., & Cleeremans, A. (2010). Measuring consciousness: is one measure better than the other? *Consciousness and Cognition*, *19*(4), 1069-1078.

Scheffers, M. K., & Coles, M. G. (2000). Performance monitoring in a confusing world: error-related brain activity, judgements of response accuracy, and types of errors. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(1), 141-151.

Schnyer, D. M., Verfaellie, M., Alexander, M. P., LaFleche, G., Nicholls, L., and Kaszniak, A. W. (2004). A role for right medial prefontal cortex in accurate feeling-of-knowing judgements: evidence from patients with lesions to frontal cortex. *Neuropsychologia, 42*, 957–966.

Sergent, C., & Dehaene, S. (2004). Is consciousness a gradual phenomenon? *Psychological Science: A Journal of the American Psychological Society/APS, 15*(11), 720–728.

Siedlecka, M., Paulewicz, B., & Wierzchoń, M. (2016). But I Was So Sure! Metacognitive judgements Are Less Accurate Given Prospectively than Retrospectively. *Frontiers in Psychology, 7* (218).

Sniezek, J. A., Paese, P. W., & Switzer III, F. S. (1990). The effect of choosing on confidence in choice. *Organizational Behavior and Human Decision Processes, 46(2)*, 264-282.

Timmermans, B., Schilbach, L., Pasquali, A., & Cleeremans, A. (2012). Higher order thoughts in action: consciousness as an unconscious re-description process. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1412-1423.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology, 26*, 1-12.

Tunney, R. J., & Shanks, D. R. (2003). Subjective measures of awareness and implicit cognition. *Memory & Cognition, 31(7)*, 1060-1071.

Van Zandt, T., & Maldonado-Molina, M. M. (2004). Response reversals in

    recognition memory. *Journal of Experimental Psychology: Learning, Memory,*

    and Cognition, *30*(6), 1147-1166.

Vickers, D., & Lee, M. D. (1998). Dynamic models of simple judgements: I.

    Properties of a self-regulating accumulator module. *Nonlinear Dynamics,*

    Psychology, and Life Sciences, *2*(3), 169-194.

Wessel, J. R., Danielmeier, C., & Ullsperger, M. (2011). Error awareness revisited:

    accumulation of multimodal evidence from central and autonomic nervous

    systems. *Journal of Cognitive Neuroscience*, *23*(10), 3021-3036.

Wierzchoń, M., Paulewicz, B., Asanowicz, D., Timmermans, B., & Cleeremans, A.

    (2014). Different subjective awareness measures demonstrate the influence of

    visual identification on perceptual awareness ratings. *Consciousness and*

    Cognition, *27*, 109-120.

Wixted, J. T., & Mickes, L. (2010). A continuous dual-process model of

    remember/know judgements. *Psychological Review*, *117*(4), 1025-1054.

Yu, S., Pleskac, T. J., & Zeigenfuse, M. D. (2015). Dynamics of postdecisional

    processing of confidence. *Journal of Experimental Psychology: General,*

    144(2), 489-510.

Zehetleitner, M., & Rausch, M. (2013). Being confident without seeing: what

    subjective measures of visual consciousness are about. *Attention, Perception, &*

    Psychophysics, *75*(7), 1406-1426.

Table 1

Average type 1 response accuracy, type 1 response RT (ms), confidence rating RT(ms) and

the average level of confidence in both conditions in Experiment 1

Condition	Accuracy	RT *type 1* response *M (SD)*	RT correct *type 1* response *M (SD)*	RT incorrect *type* 1 response *M (SD)*	RT confidence ratings *M (SD)*	Confidence for correct *type* 1 response *M (SD)*	Confidence for incorrect *type* 1 response *M (SD)*
R-CR	86%	1955 (773)	1872 (721)	2443 (886)	733 (508)	2.42 (0.84)	1.48 (0.97)
CR-R	84%	1749 (716)	1678 (668)	2112 (837)	1367 (932)	1.62 (0.92)	1.47 (0.91)

Table 2

Average confidence rating RT (ms) and the frequency for each scale point in each condition in Experiment 1

		Guessing	Not confident	Quite confident	Very confident
R-CR	**RT** *M (SD)*	766 (574)	991 (671)	864 (511)	605 (398)
	Frequency	6%	14%	25%	55%
CR-R	**RT** *M (SD)*	1559 (1004)	1459 (1038)	1289 (829)	1147 (826)
	Frequency	12%	36%	34%	18%

Table 3

Regression coefficients for the logistic regression mixed model for accuracy in Experiment 1

Number of participants: 36 Number of observations: 4281	Estimate	SE	z	*p (two-tailed)*
Intercept	-0.23	.15	-1.52	.13
Order CR-R	1.58	.22	7.06	< .001***
Confidence rating	1.10	.09	11.92	< .001***
Confidence rating: Order CR-R	-0.86	.14	-6.17	< .001***

likelihood ratio: $\chi 2(4) = 288$, $p < .001$

* $p < .05$; ** $p < .01$; *** $p < .001$

Table 4

Average type 1 response accuracy, type 1 response RT (ms), confidence RT (ms) and the average level of confidence in both conditions in Experiment 2

Condition	Accuracy	RT *type 1* **response** *M (SD)*	RT **correct** *type 1* **response** *M (SD)*	RT **incorrect** *type 1* **response** *M (SD)*	RT **confidence rating** *M (SD)*	Confidence **for correct** *type 1* **response** *M (SD)*	Confidence **for incorrect** *type 1* **response** *M (SD)*
R-CR	79%	862 (495)	816 (472)	1033 (539)	738 (481)	3.26 (0.88)	2.62 (0.92)
CR-R	77%	536 (330)	519 (308)	595 (389)	1173 (606)	3.25 (0.92)	2.72 (0.95)

Table 5

*Average confidence rating RT (ms) and the frequency for each scale point in each condition
in Experiment 2*

		Guessing	Not confident	Quite confident	Very confident
R-CR	**RT** *M (SD)*	971.19 (566)	889.84 (519)	829.74 (493)	575.76 (388)
	Frequency	6%	20%	30%	44%
CR-R	**RT** *M (SD)*	1375.66 (669)	1425.46 (604)	1334.45 (590)	945.70 (521)
	Frequency	6%	21%	26%	47%

Table 6

Regression coefficients for the logistic regression mixed model for accuracy in Experiment 2

Number of participants: 55 Number of observations: 6493	Estimate	SE	*z*	*p (two-tailed)*
Intercept	-0.10	0.12	-0.88	.38
Order CR-R	0.18	0.14	1.30	.20
Confidence rating	0.79	0.07	11.76	<.001***
Confidence rating: Order CR-R	-0.17	0.09	-1.90	.03 (*one tailed*)*

likelihood ratio: $\chi2(5) = 430$, $p < .001$

* $p < .05$; ** $p < .01$; *** $p < .001$

Figure 1. Procedure of the task in Experiment 1 (A) and Experiment 2 (B). Conditions: R-CR: response followed by confidence rating, CR-R: confidence rating followed by response.

Figure 2. Average accuracy of *type 1* responses, scale response frequency and model fit for the relationship between response accuracy and confidence ratings in each condition in Experiment 1 (R-CR, response-confidence rating; CR-R, confidence rating-response). The position of filled circles represents average accuracy for each scale point. The frequency describes the proportion of each confidence rating response in each condition.
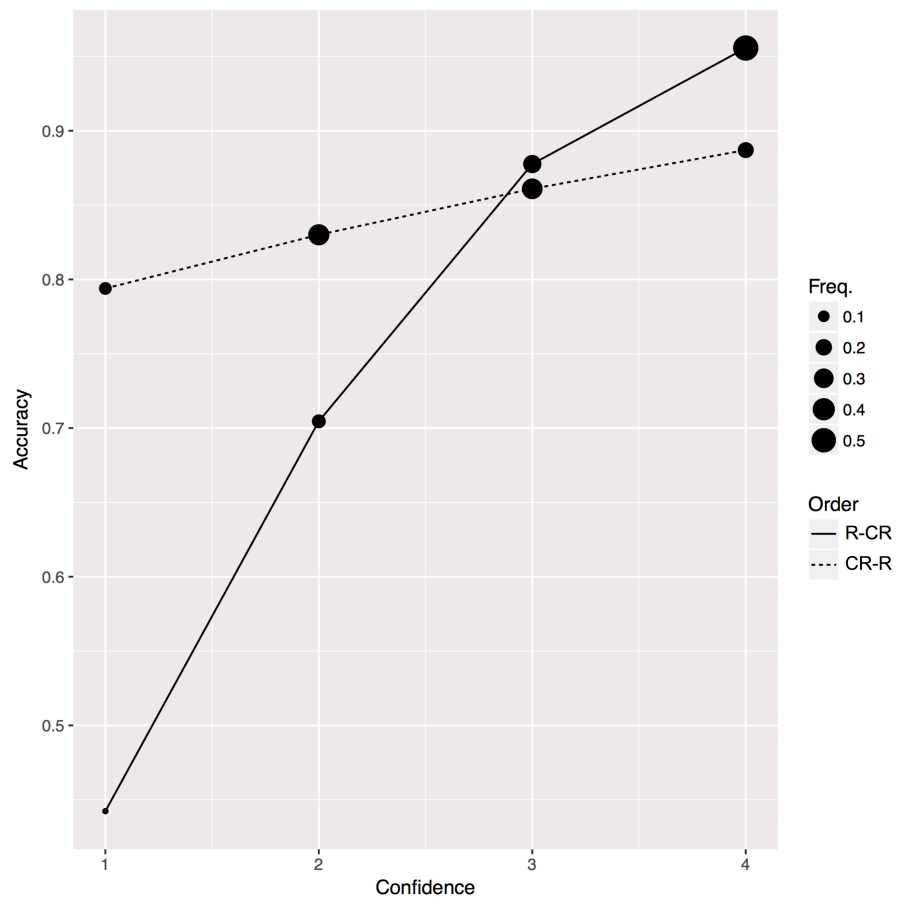
Figure 3. Average accuracy of *type 1* responses, scale response frequency and model fit for the relationship between accuracy and confidence ratings in each condition in Experiment 2 (R-CR, response-confidence rating; CR-R, confidence ratind-response). The position of filled circles represents average accuracy for each scale point. The frequency describes the proportion of each confidence rating response in each condition.