Impulsivity and Overall Similarity Classification

Andy J. Wills (andy.wills@plymouth.ac.uk), Chris A. Longmore
Department of Psychology, Plymouth University, UK.

Fraser Milton
Department of Psychology, Exeter University, UK.

Abstract

It is sometimes argued that implementation of an overall similarity classification is less effortful than implementation of a single-dimension classification. One piece of evidence taken to be in support of this argument is that highly impulsive individuals appear to be more likely to sort on the basis of overall similarity than individuals with low impulsivity (Ward, 1983); presumably, higher impulsivity results in lower effort. In the current article, we identify some limitations in Ward's procedure and, using a more standard measure of impulsivity and a less ambiguous measure of overall similarity classification, re-investigate the relationship between impulsivity and overall similarity classification. Using a match-to-standard procedure, the current experiment finds that overall similarity classification is *less* prevalent in highly impulsive individuals. The implications of this result, which is opposite to that reported by Ward (1983), are discussed.

Keywords: impulsivity; categorization; overall similarity; family resemblance.

In a seminal article, Brooks (1978) argued for two different processes of categorization. In analytic categorization, the participant separates aspects of the stimulus and evaluates their ability to predict category membership. This process of analysis, Brooks assumed, will typically lead to a subset of the stimulus attributes controlling responding. In contrast, nonanalytic categorization is the process of predicting category membership on the basis of overall similarity to known examples—a process that results in all stimulus attributes having some control over responding. Brooks hypothesized that nonanalytic categorization would be more likely to occur where cognitive resources were limited.

Brooks's hypothesis is striking because it assumes that a categorization process that employs all the information in the stimulus (overall similarity) is less effortful than a categorization process that employs a subset of that information (analytic, or "rule-based", categorization). Following Wills, Milton, Longmore, Hester, and Robinson (2013), we describe this as the *less-is-more* hypothesis—for example, less time spent categorizing objects results in more information from those objects having control over responding (Smith & Kemler Nelson, 1984). We contrast this with the *more-is-more* hypothesis—for example, more time spent categorizing objects results in more information from those objects having control over responding (Milton, Longmore, & Wills, 2008).

The current paper revisits one particular plank in the *less-is-more* argument; namely, the result reported by Ward (1983) that highly impulsive individuals are more likely to classify on the basis of overall similarity than individuals with low impulsivity. This result appears to support the less-is-more hypothesis because, presumably, impulsive individuals devote fewer cognitive resources to the categorization task than do reflective individuals.

We had two concerns about Ward's demonstration—the validity of the measure of impulsivity, and the validity of the measure of overall similarity classification. Below, we outline those concerns, and describe how we addressed them in the current study.

Impulsivity

Ward used the Matching Familiar Figures measure of impulsivity (Kagan, 1965), a measure whose validity has been questioned (e.g. Block, Block, & Harrington, 1974) and which appears to be largely uncorrelated with better validated measures of impulsivity (Helmers, Young, & Pihl, 1995). In the current study, we employed the Barratt Impulsivity Scale (BIS-11), which is the most widely used measure of impulsivity (Stanford et al., 2009). It has high reliability and good external validity (Patton, Stanford, & Barratt, 1995). The Barratt Impulsivity Scale is a self-report measure that includes statements such as, "I concentrate easily" and "I am happy-go-lucky".

Overall similarity classification

Ward employed the triad procedure as a measure of the prevalence of overall similarity classification. In this procedure three stimuli, whose relationship to each other is illustrated in Figure 1, are presented simultaneously and participants are asked to decide which two stimuli go together best. Stimuli B and C are similar on both stimulus dimensions, but not identical on either, while stimuli A and B are identical on one stimulus dimension but quite dissimilar on the other. Three responses are possible—an AB response (A and B go together best), a BC response or an AC response. Time pressure, concurrent load, impulsivity, and instructions to respond impressionistically, all increase BC responses and decrease AB responses (Smith & Kemler Nelson, 1984; Ward, 1983; Ward, Foley, & Cole, 1986). AB responding is typically described as "dimensional" responding and BC responding is typically described as "overall similarity" responding, hence leading to the claim that overall similarity (BC) responding increases as cognitive resources

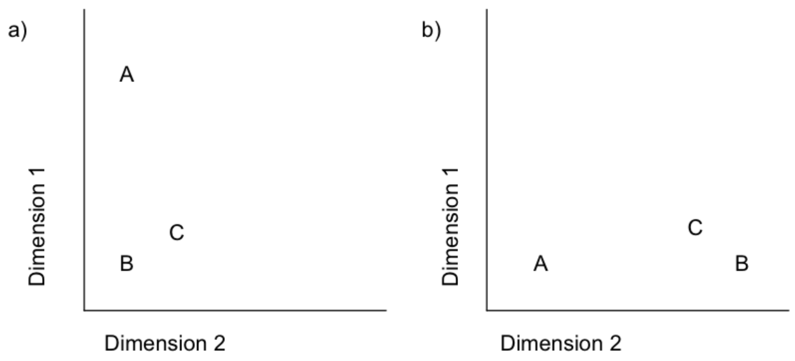

Figure 1: Abstract structure of the triad task. Typically, on half of trials A and B are identical on dimension 1 (Panel b), and on the other half of trials A and B are identical on dimension 2 (Panel a).

decrease.

One reason that the triad procedure is ill suited to testing a less-is-more hypothesis is that consistent AB ("dimensional") responding requires that the participant consider both stimulus dimensions on every trial. This is because the dimension on which A and B are identical varies unpredictably from trial to trial (see Figure 1), and so a consistent AB responder cannot decide in advance of stimulus onset to only attend to one of the stimulus dimensions. Consistent BC responding also requires consideration of both stimulus dimensions on every trial (irrespective of whether one believes that consideration to take the form of an analytic strategy or direct access to similarity relations through holistic "blobs"). Hence both consistent "overall similarity" (BC) responding and consistent "dimensional" (AB) responding requires consideration of all the relevant stimulus information on every trial. It is therefore not the case that overall similarity responding requires consideration of more of the available stimulus information than dimensional responding in the triad task, and hence the triad task is not well suited to testing a less-is-more hypothesis.

In the current study, we employed the match-to-standards procedure, which is perhaps best considered as an interpretatively less ambiguous version of the triad procedure. The procedure was introduced by Regehr and Brooks (1995) as a means of increasing the prevalence of overall similarity classification of novel stimuli, relative to the more commonly employed array sort procedure, in which single-dimension classification dominates (Medin, Wattenmaker, & Hampson, 1987). In line with the findings of Milton et al. (2008), Regehr and Brooks observed that reaction times were longer for overall similarity classification than single-dimension classification in this procedure, although they did not publish these observations (Brooks, personal communication, 20 October 2009).

In the match-to-standards procedure, participants sequentially free classify each of a series of target stimuli as belonging to one of two categories. The two categories are represented by two standards—that is, two stimuli that appear on each trial. The two standards differ from each other on all variable stimulus dimensions. For example, in the current study, the two standards are as shown in Figure 2. In the current experiments, there are 10 distinct to-be-classified stimuli, with the abstract structure shown in Table 1. In some respects, the match-to-standards procedure is similar to the triad procedure, because each trial involves deciding which two of the three stimuli go together best (although, unlike the triad task, the option of saying that the two least similar stimuli—the two standards in the match-to-standards procedure—go together is not available). Also, in the match-to-standards procedure, each participant's classification strategy for a particular block is determined over 10 trials (rather than independently for each trial, as in the triad task).

Experiment

In summary, the current experiment re-investigates the relationship between impulsivity and overall similarity classification, first reported by Ward (1983), but using improved measures of both impulsivity and of overall similarity classification.

Method

Participants and apparatus Thirty-six participants from the University of Exeter took part in the experiment in return for course credit or payment. The stimuli were presented on a 17-inch CRT monitor, set to a resolution of 800 × 600 pixels and a color depth of 16-bits per pixel. The participants sat in front of the computer screen at a distance of approximately 50 cms. Responses were made using a standard keyboard.

Stimuli The abstract stimulus structure can be seen in Table 1. The stimulus set consisted of four binary-valued dimensions (D1-D4) and the stimuli were organized around two prototypes, each representative of the two categories. These prototypes were constructed by taking all the positive values on the dimensions for one of the stimuli (1, 1, 1, 1) and all of the zero values on the dimensions (0, 0, 0, 0) for the other category. The rest of the stimuli (one-aways) were mild distortions of the two prototypes in that they had three features characteristic of their category and one atypical feature more characteristic of the other category. In total there were 10 stimuli in the set. Sorting the stimuli by overall similarity, as shown in Table 1, maximizes within-group similarities and minimizes between-group similarities. The stimuli were one of the lamp stimulus sets used by Milton and Wills (2004). Each lamp had four variable features; lampshade (with either 5 or 10 dots), width of stand (wide or narrow), color of bar (light or dark blue) and size of base (long or short). See Figure 2 for the prototypes of each category.

Table 1: Abstract stimulus set

Category A				Category B			
D1	D2	D3	D4	D1	D2	D3	D4
1	1	1	1	0	0	0	0
1	1	1	0	0	0	0	1
1	1	0	1	0	0	1	0
1	0	1	1	0	1	0	0
0	1	1	1	1	0	0	0

Procedure We used the match-to-standards task introduced by Regehr and Brooks (1995), and developed by our lab (Milton & Wills, 2004; Milton et al., 2008; Wills et al., 2013). Each trial began with a blank screen that was presented for 1000 ms, followed by three lamps in a triangular arrangement. There were two lamps at the top of the screen that depicted the prototypes for category A and B with the to-be-classified lamp presented below the prototypes. Each stimulus array remained on the screen until the participant placed the to-be-classified lamp into either category A (by pressing the 'c' key on the keyboard) or category B (by pressing the 'm' key). After the participant made a response the next trial began.

There was no feedback; participants were simply instructed at the beginning of the experiment to sort the stimuli in the way they thought most appropriate. At the end of each block, participants were asked to write down the sorting strategy they used before moving onto the next block. Participants were presented with a total of 60 trials, in 6 blocks of 10 trials. In each block, each of the stimuli shown in Table 1 was presented once as the to-be-classified stimulus. The order of presentation within a block was random.

Immediately after the 6 blocks of classifiation, participants' impulsivity was assessed using the Barratt Impulsivity Scale BIS-11 (Stanford & Barratt, 1995).

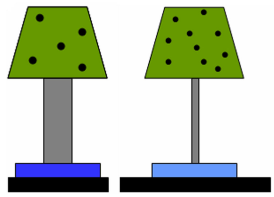

Figure 2: Stimulus prototypes

Classification measure Each participant was classified as having produced one of the sort types described below in each of the six blocks of the experiment. These sort types are identical to those used previously by our lab (Milton & Wills, 2004; Milton et al., 2008; Milton, Wills, & Hodgson, 2009; Wills et al., 2013).

A *uni-dimensional sort* is based on a single dimension of the stimulus. It does not matter which dimension is used as the basis of sorting, so long as all of the positive values for the chosen dimension were in one category and all of the zero values for that dimension were in the other category. Additionally, in order to receive this classification, the participant has to describe their sort as being based on a single dimension.

Participants were considered to have produced a *one-away uni-dimensional sort* if they described their sorting as being driven by a single dimension but there was a solitary error in their classification. This means that nine of the items were classified on the basis of a single dimension but the other item was placed into the wrong category.

An *overall similarity sort*, also commonly known as a "family resemblance" sort (Medin et al., 1987), has a structure identical to that shown in Table 1. In order to receive this classification, the participant had to place each of the prototypes, along with their derived one-aways, into separate categories without error. Additionally, they have to describe their strategy as being based either on general similarity or on placing each item into the category with which it had more features in common.

A *one-away overall similarity sort* is similar to the one-away uni-dimensional sort with the exception that the error occurred in a sort that was otherwise overall similarity.

Any classifications produced by a participant other than those described above were classified as other sorts, even if the description given by the participant fitted one of the sorts described above. The correspondence between the classification produced by a participant and their verbal description of the sort they have produced is very high in this procedure, approximately 0.99. The verbal descriptions were classified by the authors.

Impulsivity measure Participants were classified as *high impulsivity* if their score on BIS-11 was greater than the sample median (64.5), and as *low impulsivity* otherwise (cf. Martin & Potts, 2009).

Results

For every block, each participant's sorting strategy was classified according to the sort types described above. As in previous studies, one-away uni-dimensional and one-away overall similarity sorts were classified as uni-dimensional and overall similarity sorts respectively (Milton & Wills, 2004; Milton et al., 2008; Wills et al., 2013). The total number of sorts for each strategy was calculated and the mean proportions of overall similarity, uni-dimensional and other sorts produced by high impulsivity and low impulsivity participants are shown in Figure 3.

Participants with high impulsivity produced significantly more uni-dimensional sorts than those with low impulsivity, $t(34) = 2.203, p < .05$. Conversely, participants with high impulsivity produced significantly fewer overall similarity sorts than those with low impulsivity, $t(34) = -2.382, p < .025$. There was no difference in the prevalence of Other sorts, $t(34) = .206, p = .838$. Correlations of the raw impulsivity scores with the prevalence of unidimensional, overall similarity and other sorts reveal the same ordinal pattern, albeit with slightly higher p-values (unidimensional $\tau = .22, p = .09$; overall similarity $\tau = -.20, p = .13$; other, $\tau = -.06, p = .68$).

Discussion

Ward (1983) reported that participants with high impulsivity were more likely to classify on the basis of overall similarity than participants with low impulsivity. This is one of a number of results taken to support the idea that overall similarity classification is a low-effort, "fall-back" mode of classification that people employ when cognitive resources are limited. However, a close examination of Ward's study reveals that both the measure of impulsivity, and the measure of overall similarity, employed are sub-optimal. The measure of impulsivity (Kagan's Matching Familiar Figures task) is of questionable validity, and does not correlate with other more valid measures of impulsivity. The triad task, employed by Ward as a measure of overall similarity responding, is also interpretatively ambiguous because both consistent overall similarity responding, and consistent dimensional responding, require consideration of both stimulus dimensions on every trial.

In the current study, we employed a more standard measure of impulsivity (the Barratt Impulsivity Scale), and measured overall similarity responding with the match-to-standards task. The match-to-standards task is a variant of the triad procedure that overcomes the interpretive ambiguities in the standard procedure. One of the ways it achieves this is by considering the participants' responses to a series of ten stimuli, rather than considering the response to each stimulus as an independent data point.

Our refinement of Ward's procedures seems to have led to a reversal of his conclusions. In the current study, high impulsivity is associated with single-dimension responding, whilst low impulsivity is associated with overall similarity responding. Thus, our data seem to support a conclusion opposite to Ward's—overall similarity classification is more requiring of cognitive resources than dimensional responding. Such a conclusion is consistent with previous results employing the match-to-standards procedure. For example, Milton et al. (2008) found that time pressure generally reduces the prevalence of overall similarity responding, Milton et al. (2009) found greater frontal lobe involvement for overall similarity classification than single-dimension classification, and Wills et al. (2013) found that concurrent load, and a small working memory capacity, reduces the prevalence of overall similarity classification, and that instructions to respond meticulously increased overall similarity responding. Milton and Wills (2009) found that overall similarity classification takes longer, and involves more, widely distributed, eye movements than single-dimension classification. Taken together with the results of the current study, a consistent picture is emerging—overall similarity classification is more effortful than single-dimension classification.

One key question, not satisfactorily answered by this study, or by any other published study, is whether the consistent pattern of results emerging from the match-to-standards procedure is specific to that procedure. Perhaps the match-to-standards procedure is the exception, with other procedures pointing consistently to the opposite conclusion? In our view, there is currently insufficient data to answer this question adequately. The triad procedure, at least as typically analysed, is interpretatively ambiguous, but this problem could be overcome with larger samples and more sophisticated model-based analyses (e.g. Thompson, 1994).

Two other procedures that are sometimes taken to support the less-is-more view (that overall similarity classification is lower effort than single-dimension classification) are the Ashby-Maddox procedure (Ashby & Maddox, 2005) and the criterial-attribute procedure (Kemler Nelson, 1984; Smith & Shapiro, 1989). As discussed by Wills et al. (2013), these procedures seem likely to be addressing slightly different questions to the one posed here.

The Ashby-Maddox procedure seems, predominately, to be an investigation of the effects of having an easy-to-verbalize category structure versus a hard-to-verbalize structure. As single-dimension structures are typically easy to verbalize, and some multi-dimensional structures are not, the two issues are not unrelated. However, recent work by Ashby, Maddox and colleagues sug-

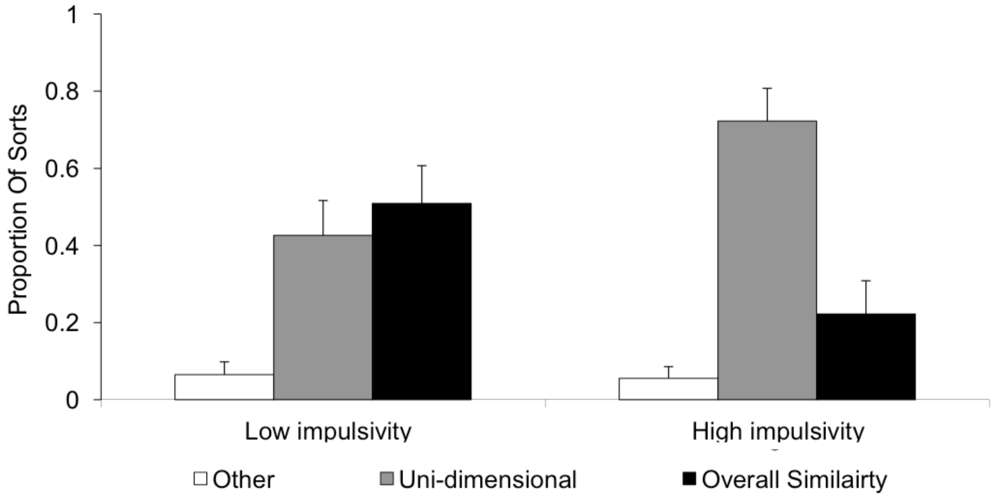

Figure 3: Proportion of other, uni-dimensional and overall similarity sorts, by impulsivity. Error bars represent one standard error

gests that it is verbalizability, rather than dimensionality, that underlies their reported effects, because the effects are still observed when one compares two multi-dimensional classification problems that differ in verbalizability (Filoteo, Lauritzen, & Maddox, 2010; Maddox, Pacheco, Reeves, Zhu, & Schnyer, 2010). There are also an increasing number of studies that suggest that some of the results from the Ashby-Maddox procedure are a consequence of subtle problems with the design or analysis of these studies (Newell, Dunn, & Kalish, 2010; Newell, Moore, Wills, & Milton, 2013).

The criterial-attribute procedure, like the Ashby-Maddox procedure, seems to be addressing a slightly different question to the one posed here. Specifically, the criterial-attribute procedure may provide evidence that it is effortful to detect the one dimension that permits above-criterion performance in a context where all dimensions individually support at least 75% accuracy. Studies using the criterial-attribute procedure (Kemler Nelson, 1984; Smith & Shapiro, 1989) support the idea that this is effortful, but they are also consistent with the idea that *implementing* an overall similarity classification is more effortful than implementing a single-dimension classification (Smith, Tracy, & Murray, 1993). It was the implementation of overall similarity classification that was the topic of the current study.

Why did the current study and Ward's study produce apparently opposite results regarding impulsivity and overall similarity classification? There are a number of possibilities. One possibility, as previously discussed, is that Ward's measure of impulsivity has low validity. If Ward did not measure impulsivity adequately, then contrasting the current result (using a more valid measure) with Ward's findings is largely irrelevant, as Ward's results would not validly concern impulsivity. Another possibility is that a less ambiguous analysis of the triad task using model-based methods might reveal that high impulsivity was in fact associated with uni-dimensional classification, rather than overall similarity classification, even in the triad task. We are currently investigating this possibility.

One way in which the match-to-standards task differs from other procedures is that the participants' written descriptions of their sorts are combined with the sorts they actually produce in order to classify their behavior. It seems likely that impulsivity affects the content of those written reports. A more critical possibility is that impulsivity might affect the written reports differently to the objective sorts, and hence the results of the current study might have been different if we had only looked at the objective sorts (or only looked at the written reports). However, if impulsivity does affect written reports differently to the objective sorts, then the consequence should be a difference in the prevalence of Other sorts as a function of impulsivity (because Other sorts occur under our classification procedure when the written report and objective sort do not agree). As can be seen in Figure 3, the proportion of Other sorts is low and does not vary by impulsivity. It therefore seems unlikely that our results would have been substantively different if we had considered just the objective sorts or just the written descriptions.

In conclusion, in the current study highly impulsive individuals were more likely to produce single-dimension classifications than low-impulsivity individuals (who were more likely to produce overall similarity classifications). The opposite conclusion, suggested by Ward (1983), seems likely to be due to limitations of the procedures employed in that study, although it remains a possibility that both results are valid, but specific to the procedure employed.

Acknowledgments

This research was supported by ESRC grant RES-000-22-1779 awarded to Andy J. Wills and ESRC grant PTA-026-27-1256 and a Great Western Research Fellowship awarded to Fraser Milton.

References

Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology*, *105*, 442–481.

Block, J., Block, J., & Harrington, D. (1974). Some misgivings about the matching familiar figures test as a measure of reflection-impulsivity. *Developmental Psychology*, *10*, 611–632.

Brooks, L. R. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorization* (pp. 169–211). Hillsdale, NJ: Lawrence Erlbaum Associates.

Filoteo, J., Lauritzen, S., & Maddox, W. T. (2010). Removing the frontal lobes: The effects of engaging executive functions on perceptual category learning. *Psychological Science*, *21*, 415–423.

Helmers, K., Young, S., & Pihl, R. (1995). Assessment of measures of impulsivity in healthy male volunteers. *Personality and Individual Differences*, *19*, 927–935.

Kagan, J. (1965). Impulsive and reflective children: Significance of conceptual tempo. In J. Krumboltz (Ed.), *Learning and the educational process.* Chicago: Rand McNally.

Kemler Nelson, D. (1984). The effect of intention on what concepts are acquired. *Journal of Verbal Learning and Verbal Behavior*, *23*, 734–759.

Maddox, W. T., Pacheco, J., Reeves, M., Zhu, B., & Schnyer, D. M. (2010). Rule-based and information-integration category learning in normal aging. *Neuropsychologia*, *48*, 2998–3008.

Martin, R., & Potts, G. (2009). Impulsivity in decision making: An event-related potential investigation. *Personality and Individual Differences*, *46*, 303–308.

Medin, D. L., Wattenmaker, W. D., & Hampson, S. E. (1987). Family resemblance, conceptual cohesiveness, and category construction. *Cognitive Psychology*, *19*, 242–279.

Milton, F., Longmore, C. A., & Wills, A. J. (2008). Processes of overall similarity sorting in free classification. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 407–415.

Milton, F., & Wills, A. J. (2004). The influence of stimulus properties on category construction. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *30*, 407–415.

Milton, F., & Wills, A. J. (2009). Eye movements in overall similarity and single-dimension sorting. In *Proceedings of the thirty-first annual conference of the cognitive science society* (pp. 1512–1517). Austin, TX: Cognitive Science Society.

Milton, F., Wills, A. J., & Hodgson, T. L. (2009). The neural basis of overall similarity and single-dimension sorting. *NeuroImage*, *46*, 319–326.

Newell, B. R., Dunn, J. C., & Kalish, M. (2010). The dimensionality of perceptual category learning: a state-trace analysis. *Memory and Cognition*, *38*, 563–581.

Newell, B. R., Moore, C. P., Wills, A. J., & Milton, F. (2013). Reinstating the frontal lobes? Having more time to think improves "implicit" perceptual categorization. a comment on Filoteo, Lauritzen and Maddox, 2010. *Psychological Science*, *24*, 386–3389.

Patton, J., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt Impulsiveness Scale. *Journal of Clinical Neuroscience*, *51*, 768–774.

Regehr, G., & Brooks, L. R. (1995). Category organization in free classification: The organizing effect of an array of stimuli. *Journal of Experimental Psychology: General*, *122*, 92–114.

Smith, J. D., & Kemler Nelson, D. (1984). Overall similarity in adults' classification: The child in all of us. *Journal of Experimental Psychology: General*, *113*, 137–159.

Smith, J. D., & Shapiro, J. (1989). The occurrence of holistic categorization. *Journal of Memory and Language*, *28*, 386–399.

Smith, J. D., Tracy, J. I., & Murray, M. J. (1993). Depression and category learning. *Journal of Experimental Psychology: General*, *122*, 331-346.

Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt Impulsivity Scale. *Journal of Clinical Neuroscience*, *51*, 768–774.

Stanford, M. S., Mathias, C. W., Dougherty, D., Lake, S., Anderson, N., & Patton, J. (2009). Fifty years of the Barratt Impulsiveness Scale: An update and review. *Personality and Individual Differences*, *47*, 385–395.

Thompson, L. (1994). Dimensional strategies dominate perceptual classification. *Child Development*, *65*, 1627–1645.

Ward, T. B. (1983). Response tempo and separable-integral responding: Evidence for an integral-to-separable processing sequence in visual perception. *Journal of Experimental Psychology: Human Perception and Performance*, *9*, 103-112.

Ward, T. B., Foley, C., & Cole, J. (1986). Classifying multidimensional stimuli: Stimulus, task, and observer factors. *Journal of Experimental Psychology: Human Perception and Performance*, *12*, 211–225.

Wills, A. J., Milton, F., Longmore, C. A., Hester, S., & Robinson, J. (2013). Is overall similarity classification less effortful than single-dimension classification? *Quarterly Journal of Experimental Psychology*, *66*, 299–318.