Reviewers' comments:

## Editor (Weiwei Zhang):

Our sincere thanks for letting us consider your manuscript at Cognition. Three expert reviewers' comments on your work have now been received. You will see that they are advising against publication of your work. Although the reviewers like the research question, they have some serious concerns on the experimental design (e.g., stimulus durations conceptualized as time pressure and its expected effects on perceptual encoding, Reviewers #1&3), study rationale (e.g., conceptualization of integrality, reviewers 2&3), and statistic power (reviewer #2).  I am in agreement with these concerns and have therefore decided to reject this submission but invite a major revision. Some of the issues highlighted by the reviewers may be addressable without new data.

Thanks for the thoughtful comments of both you and the reviewers. To summarise:

* Reviewer 1 “supports publication in Cognition” and considers their comments minor. We have addressed these minor comments.
* Reviewer 2’s comments primarily involve requests for clarification and addition, along with a few misconceptions due to a lack of clarity on our part. We have acceded to those requests and clarified where needed.
* Reviewer 3 thinks we have an “interesting result” that “deserves further attention”. They have some “questions and possible reservations” which they consider may be “addressed in any requested revision”. We agree with this assessment and have done so.

I would also like to chime in on the visual persistence issue with the shorter stimulus duration condition (reviewer #3, Dr. Pomerantz). Given the lingering visual persistence after the stimulus offset for up to 500 ms, perceptual encoding would be likely completed. Would any effects reported here a result of some strategical factors?

*Yes, that’s possible. We’re confident that our results are robust, but the precise mechanisms that underlie the effect remain uncertain and Combination Theory does not make any direct predictions in this regard. Categorization involves a number of processes (e.g., perceptual, attentional, decisional/strategic and action) that are unlikely to be independent of each other, making partitioning the precise contribution of these various difficult. Nonetheless, these are interesting questions for future research.* ***These points are now considered in the General Discussion (p. 19, line 377).***  
  
Reviewer #1: This paper reports 3 experiments investigating the effect of time pressure on how participants classify integral stimuli. Using a triad task with a set of colored squares varying in chroma and brightness, the authors showed that increasing time pressure increased the use of unidimensional classification. Critically, these results are inconsistent with the idea that integral stimuli are perceived as integrated configurations that require resources for selective attention to a single dimension (Differentiation Theory). Rather, the results support the idea that integral stimuli are initially perceived as separable dimensions which are quickly combined (Combination Theory).  
  
I thought this paper was well-written, tested a clear prediction capable of distinguishing between two competing theories, and the methods and analyses were rigorous and clearly presented. The paper tests a novel, interesting, and counter-intuitive prediction, that a specific type of behavior considered effortful by Differentiation Theory will be easier when participants are given less time to respond. Although by-block classification seemed necessary to detect the effect, the effect was demonstrated in all 3 experiments and Experiment 3 showed that a MDS-informed re-analysis strengthened the effect. The results offer strong support for the authors' conclusions regarding Combination Theory, and have more general implications for perception and cognition (discussed briefly). I enjoyed reading this paper and support its publication in Cognition. I have only minor comments for the authors to consider.

*Thank you, we are glad you enjoyed it.*   
  
Could report proportion in addition to frequency in Table 1 to facilitate comparison between later tables.

*Sure. Added as requested.* ***See Table 1, page 10.***  
  
Final sentence in first paragraph: should this say "and are considered as primary or privileged stimulus dimensions"?

*Yes, this has been edited –* ***see Line 41-43, page 3****.*   
  
I thought the term "p-hacking" on page 11 was a little strong. I think it's sufficient to describe the analysis in Experiment 1 as post-hoc, especially since the result replicates.

*We have taken your advice and have removed the reference to p-hacking and instead made it clear that it was a post-hoc analysis –* ***see Lines 229-231, page 12.***   
  
If time pressure allows participants to selectively attend to each dimension prior to combination, this should make the integral stimuli more like separable stimuli. The Introduction states that for separable stimuli, there is a strong preference for identity responses, but this didn't occur in the current study. Experiment 3 shows that this might be because the stimuli were not perceived as exactly identical on the dimensions, however this could also be true for studies demonstrating ID responses for separable stimuli. The authors could add some discussion for why ID responses were not more prevalent. For example, is it that UD responding is easier since it doesn't require switching attention between dimensions? If so, why doesn't UD responding also dominate classification of separable stimuli? Or, maybe it does, but the Introduction made it sound like OS was the dominant strategy. The point here is that there still seems to be some difference between separable, and (pre-combination) integral stimuli, which the authors might want to discuss.  
  
*As requested, we have added a discussion of the low prevalence of ID responses in our experiments, relative to experiments with separable stimuli -* ***General Discussion, p. 18, starting line 345****.*

Reviewer #2:  
In this paper, the authors concluded that integral stimuli are not slowly analyzed, but quickly synthesized. The authors would like to use the Differentiation Theory to explain this phenomenon.

- The Combination Theory assumes that the dimensions of integral stimuli are separable at the first place, however, they are processed holistically, and that's why they are called "integral stimuli". If this is not the intent of what the authors are trying to say in the Introduction section, then the authors should explain this issue more clearly.

*This summary doesn't quite align with the message that we wished to convey. We compared the predictions of Differentiation Theory and Combination Theory in explaining patterns of categorisation behaviour, and came out in favour of Combination Theory (not Differentiation Theory). Combination Theory argues that cognition begins with the stimulus attributes and that they are combined if time and/or mental resources allow. Combination Theory does not assume that integral stimuli are processed holistically.* ***We’ve edited the first five paragraphs of our Introduction to make this clearer.***

- Please explain more about the "noisy" in the sentence "A strong preference for overall similarity classification in integral stimuli under low time pressure is already well established, and the application of time pressure makes classification data more 'noisy'".

*Sure, now done -* ***footnote, p. 7***

- Please add information about the IRB number.

*Happy to add –* ***this has now been added to the Author Note on the title page.***

- In the power analysis, the statistical analysis method should be mentioned.

*This has now been added – please see* ***line 145-151, page 8 and line 235-237, page 12.***

- It is hard to understand "There are six different ways in which three stimuli can be placed in three spatial locations and thus each of the eight triads had six different instantiations, resulting in 48 physically different triads per experiment". Please provide the readers with examples in case that the spatial structure of objects formed in the mind of readers is inconsistent with the authors' intent.

*Thanks, good suggestion. We’ve added examples – see* ***line 160, page 8****.*

- Please be more specific about how the performance of participants was classified into UD, OS, ID, or Bias, and what are the meanings of these four words.

*Sure. We’ve added a “Strategy analysis” section to the method to achieve this –* ***see page 9-10****.*

- Why an experiment with the same stimuli, apparatus, and procedure as Experiment 1 was conducted?

*This experiment was conducted as the key conclusions of the initial experiment were based on post-hoc analyses. This is now clarified in* ***lines 229-231, page 12****.*

- The authors mentioned "Crucially, increased time pressure once again increased the prevalence of unidimensional classification of these integral stimuli, BF10= 1047." What are the two groups that are comparing?

*The longer (2000ms) and shorter (100 ms) presentation-time groups. We now clarify this on* ***line 243, page 12****.*

- The sample size of Experiment 3 is limited.

*The sample size for Experiment 3 is in line with past MDS experiments in categorization research (examples now given in the manuscript) and is consistent with the sort of sample size shown by Rodgers (1991, Applied Psychological Medicine) to lead to good metric recovery. We now clarify this on* ***line 264-267, page 13****.*

- Please give some examples of the procedure of Experiment 3.

*We’ve added a Figure illustrating an example trial in Experiment 3 –* ***see page 14****.*

- The conclusions drawn in this study are based on two dimensional integral stimuli. However, if more dimensions are involved, different conclusions might be found.

*We agree that it remains an open question as to what the pattern of results would be with three dimensional integral stimuli.* ***We now note this in the General Discussion (page 21, line 411-412)*** *as a fruitful line for further research.*

- In the Vigo et al. (2022) paper, a Dual Discrimination Invariance Model was proposed to account for the behavioral results of the integral stimuli. Please refer to it and check if it helpful to explain the results observed in this study and the discussion between Euclidean or Manhattan distance in psychological space.

DDIM is an interesting model of classification, and we are happy to **include some discussion of it alongside other models in our General Discussion – see p. 22.**

Reviewer #3: Review of Cognition Ms. COGNIT-D-23-00794  
The rapid synthesis of integral stimuli, by Lori Holt  
Reviewed by James Pomerantz, September 2023  
  
This manuscript reports three experiments testing whether differentiation theory or combination theory better explains result from similarity judgment experiments using squares varying in brightness and saturation, which have been shown to behave as integral dimensions in experiments dating back to the 1960s. The results show support for combination theory, suggesting that integral stimulus dimensions are not initiated fused and then separated through a slow secondary process; rather, it's argued, integral dimensions start out of separate but are then fused through a quick secondary process.  
  
This is an interesting result and one that deserves further attention. I do have some questions and possible reservations about this manuscript, but they may well be addressed in any requested revision.  
  
*Thank you, we are glad you found it interesting.*

First, the primary manipulation in the experiments is stimulus exposure time, either 100ms or 2,000 ms. This is referred to as "time pressure", a term that normally refers to response time, i.e., how quickly Ss try to respond. Many of the tasks used to define stimulus integrality are speeded classification tasks, where Ss are encouraged to respond as quickly as possible, but that was not the case here - there apparently was no deadline for responding.

*Thanks for raising this. We agree that our experiments are more properly described as manipulating stimulus presentation time. As noted* ***in the Introduction (p. 4),*** *the predictions of Combination and Differentiation Theory with regards to separable stimuli have previously been compared with both stimulus presentation time procedures and response deadline (‘time pressure’) procedures. The decision to use stimulus presentation time in the current experiments was motivated by our previous use of this procedure in a comparable experiment with separable stimuli (Wills et al., 2015, Exp. 2).* ***We now make this point on p. 7, line 131-134.***

If the combination theory/hypothesis were correct, then selective attention should be easy with all stimuli, since it claims that all dimensions are initially represented separately. Yet Ss cannot (or at least do not) selectively attend to integral dimensions in the speeded classification task.

*Combination Theory does indeed state that integral dimensions are initially represented separately. However, the theory further states that integral dimensions are rapidly combined and, once combined, are hard to selectively attend. So, observing any unidimensional classification of integral stimuli will likely depend on having very short stimulus presentation times, and a highly sensitive analysis method.*

*The current experiments concern robust but subtle effects at very short presentation times. For example, in Experiment 2, a 100 ms presentation time led to an average of 14% of response blocks exhibiting unidimensional classification (up from 1% at 2000ms). This of course means that, even at this very short presentation time, classification was typically (85%) overall similarity. The theoretical point is that the direction of the effect (increased unidimensional responding with decreased stimulus presentation time) is as predicted by Combination Theory, and counter to the predictions of Differentiation Theory.* ***We now clarify this on p. 19, line 357-369.***

Also, note that while the shorter exposure duration was only 100ms, Ss would have had longer time to respond simply because of visual persistence (ala Sperling). If a mask followed the stimuli, that might change things, but no mask was used (unless the text that followed the stimuli overprinted the stimuli themselves).

*Agreed - the addition of a mask would likely have reduced the effective stimulus presentation time, for the reasons you outline. However, it seems unlikely that it would have reversed the direction of the effect of presentation time, and it is this direction that is the subject of the hypotheses under test.* ***This is now mentioned in the Discussion p. 21, line 413-418.***

Relatedly, a very short stimulus presentation may have altered the perception of the colors themselves, given that cones require more energy summed over time to be activated; if so, this would seem to complicate the interpretation of the results, so some control tasks might be needed to determine if the longer and shorter duration color patches appeared the same.

*Time to peak response is approximately 60ms for foveal cones (Masland, 2017), against an effective stimulus presentation time in our experiments of at least 100ms. So, insufficient time for the cones to respond to the stimulus is unlikely to be a sufficient explanation for the effect we observe.*

*Nonetheless, psychophysics and visual neuroscience are relevant to discussion of Combination Theory versus Differentiation Theory. For example, evidence suggests that luminance is represented at least somewhat separately from hue in the LGN (Ghodrati et al., 2017), and later combined through recurrent inhibitory activity in early visual cortex (Xing et al. 2015). The later combination of stimulus components that are initially represented separately seems more consistent with Combination Theory than Differentiation Theory.*

***We now cover these issues in the General Discussion, on p. 21, line 419-434.***

Second, the concept of integrality, as Garner conceived it, was defined by converging operations. That means that the concept was independent of any single system of measurement such as similarity metrics or free classification groupings. A second metric used was what is now called Garner Interference (Pomerantz, 1983 JEP:General; Algom, 2016), which refers to the Orthogonal condition being performed more slowly that the Control condition in Garner's paradigm. A third metric was redundancy gains, which refer to the condition with Correlated dimensions being performed more quickly than the Control condition. There were further measures, but the point is that the notion of dimensional integrality is built on a broad foundation, and any re-evaluation of that notion needs to assess more than just the one metric considered here, namely similarity.  
  
Relatedly, even Garner interference, which is probably the most widely known diagnostic for dimensional integrality, can have multiple explanations. One of them is the one characterized here as differentiation theory holds that these dimensions are initially fused and required slow, secondary processing to separate. Thus, selective attention fails. But there are other outcomes and explanations. For example, consider the classic Garner paradigm with four stimuli resulting from the crossing of two binary dimension, but where the dimensions are the direction of curvature of the left and right elements of parenthesis pairs: ((, (), )(, and )). These stimuli show high levels of Garner Interference (Pomerantz and Garner, 1973), but not because of any failure of selective attention. Rather, Ss do not process the individual parentheses but rather choose to process the pairs of parentheses, because of the emergent features they possess (most obvious is their bilateral symmetry about the vertical axis. So although this manuscript presents a result that might prove to be a challenge for differentiation theory, it will take more work to establish that Garner's concept of integrality got things backwards. As the author also points out, it will be important to extend the research here to include more examples of integral dimensions than the single one tested here (brightness and saturation).

*We thank the reviewer for their thoughts, which we agree with. In particular, we also think that there is something fundamentally different between integral and separable stimuli and that the evidence for this, as the reviewer outlines, is built on broad and robust foundations. Our work is not intended to question Garner's general concept of integrality but instead to provide an intriguing challenge to Differentiation Theory and broader support for the principles of Combination Theory.* ***We have made this clearer in the revision – see p. 19, line 377-382.***

On p. 9, a bit more needs to be said the four models (unidimensional, overall similarity, identity, and bias). I get their general nature but don't see much in the way of specifics.

*We’ve now added a strategy analysis section to the method of Experiment 1 to address this (****page 9-10)****) and have detailed the approach more carefully in the ‘The Current Experiments’ section –* ***(page 5-7)***