of specific stimulus features and responses.

Given that the feature integration account highlights potential associative confounds in the CSE that seem to be specifically inherent to small stimulus and response sets, a natural response to these concerns was a movement toward employing conflict tasks with larger sets (typically, moving from 2-AFC to 4-AFC schemes), such that first-order repetitions of stimulus and response features could be either prophylactically prevented from occurring (e.g., Puccioni and Vallesi, 2012; Jimenez and Mendez, 2013), or removed from analysis after the fact (e.g., Ullsperger et al., 2005; Akcay and Hazeltine, 2007). However, as recently highlighted by several authors (Schmidt and De Houwer, 2011; Mordkoff, 2012), this trend may have introduced a new associative confound to the CSE, in the form of *contingency learning*. Specifically, the expansion of the stimulus set (for instance, going from two to four colors in the Stroop task) creates more possible unique incongruent than congruent stimuli. When researchers then present congruent and incongruent trials with the same frequency (i.e., 50%), each congruent stimulus occurs more frequently (and well above chance) than each incongruent stimulus, which creates a contingency linking each distracter to their congruent response (e.g., the word RED is most frequently paired with the color red, and thus, the response "red"). Since high-contingency (congruent) trials are responded to faster than low-contingency (incongruent) trials, *and* consecutive trials with the same contingency level appear to facilitate performance (Schmidt and De Houwer, 2011), it is possible that the CSE in typical 4-AFC tasks is a reflection of contingency-learning rather than of control-based processing adjustments (**Figure 1C**).

At this point, it is worth to already highlight an overarch-

- What is the definition of an "associative confound?"; My understand is that having only 2 AFCs means that the associative explanation of CSE is not generalisable; to make it generalisable, a greater data set e.g. 4 AFCs is required.
- I don't understand the point they are making about contingency learning being a confound?

Commented [CB1]: Generally a confound refers to a variable that you want to control for, but was not and covaries with your variable of interest. (So it's something that prevents you from making the conclusion you'd like, because your results could have occurred due to the confound & not what you hypothesized.)

Associative cofound is a confound that is resulting from any associative learning perspective. In this particular section, Tobias highlights the perspectives of folks who don't really believe cognitive control exists – and any effects therein are really just a by product of us automatically associating particular stimuli with responses in the world. (Or at least saying more "control" effects are related to this kind of associative learning than most people say.)

So, in the first sentence, Tobias here says that in an attempt to control for associative confounds, people started using 4 AFC tasks. This is so that they could remove any "direct stimulus repetitions" (e.g., RED printed in RED twice in a row) or partial repetitions (e.g., RED printed in BLUE, BLUE printed in GREEN presented in a row). 'Prophylactically prevented from occurring' I assume refers to directly manipulating the sequence of stimuli that you present in a trial, via your code.

Commented [CB2]: For contingency learning, look at a full set of stroop stimuli… and say you want every stimulus to be presented equally (because you don't want it to be that someone is just better at categorizing one stimulus – i.e., just a general stimulus confound)

Red printed in red – 25%
Blue printed in blue – 25%
Green printed in green – 25% of congruent trials
Yellow printed in yellow – 25%

Red printed in blue – 8.3% of incongruent trials
Red printed in green – 8.3% of incongruent trials
Red printed in yellow – 8.3% of incongruent trials
Blue printed in red – 8.3% of incongruent trials
Blue printed in green – 8.3% of incongruent trials
Blue printed in yellow – 8.3% of incongruent trials
Green printed in red – 8.3% of incongruent trials
Green printed in blue – 8.3% of incongruent trials
Green printed in yellow– 8.3% of incongruent trials
Yellow printed in red – 8.3% of incongruent trials
Yellow printed in blue – 8.3% of incongruent trials
Yellow printed in green – 8.3% of incongruent trials

Now if you're also manipulating "proportion congruent" so that say one location on the screen is 80% more often associated with congruent vs. incongruent stimuli, while the bottom of the screen is 80% more often associated with incongruent vs. congruent stimuli.

There are 4 congruent trial types. 12 incongruent trial types. For a congruent location, where 80% of the time, they're congruent stimuli, that also means that 80% of the time, the person can just read the word and know what response to give. The %age for incongruent stimuli generally speaking in that scenario is less than the percentage of each individual congruent stimulus. Do you see how that'd be a problem? ...

and response features over successive trials can profoundly affect performance (Hommel, 1998) and it is *impossible* to circumvent the confounding factor of differential feature overlap between different congruency sequences in the CSE when employing small stimulus sets (e.g., only two or three different target and distracter stimuli). It is therefore possible, or even likely, that CSEs observed in studies with such small stimulus sets are partly, predominantly,

- p. 4
- Why is this a "confounding" factor in looking at CSE effects
- And bigger data sets, there are contingency learning confounds:

Second, while the movement toward employing larger stimulus sets has resulted in a number of studies reporting CSEs in the absence of feature repetitions (e.g., Ullsperger et al., 2005; Akcay and Hazeltine, 2007, 2011; Hazeltine et al., 2011), almost all of these studies appear to be open to alternative interpretation based on possible contingency-learning confounds because of above-chance occurrence of congruent stimuli (see Schmidt, 2013). Similarly, 2-AFC studies that require subjects to categorize large sets of unique stimuli (e.g., classifying face stimuli according

- P. 4
- Feature integration and contingency learning confounds?
- Overcome both these confounds:

Weissman et al., 2014; but see Mayr et al., 2003; Jimenez and Mendez, 2013). A typical design of this recent wave of studies circumvents both stimulus and response feature repetitions, as well as contingency-learning confounds, by splitting a 4-AFC task into two alternating 2-AFC tasks with non-overlapping stimulus and response sets (e.g., presenting alternately Stroop stimuli that are made up either of red/blue or of green/yellow combinations; e.g., Schmidt and Weissman, 2014). This approach has produced robust evidence for the basic presence of a "memory confound-

Commented [CB3]: If you have
Red in red
Green in green

Green in red
Red in green

And the idea of event files are that you're creating a sort of temporary memory of the target "red print color", "red color-word", "response button z" etc., and you want to avoid "feature overlap," it's nearly impossible in this case because your stimulus set is so small that on basically every trial you have at least one feature that overlapped with the previous trial (e.g., red in red followed by green in red – same print color on back to back trials; red in red followed by red in green – same color-word on back to back trials, etc.).

Commented [CB4]: Are you asking why?

'because of above-chance occurrence of congruent stimuli' – you can see that in the example above

Commented [CB5]: I'm not sure I know what the question is here?

Commented [CB6]: Also can see that in the example I gave above with the frequency biased vs. unbiased stimuli. You'd basically change the biased stimuli for the different contexts you're trying to create and look at these nonpredictive stimuli to assess what participants are really doing.