**DATA CLEANING**

Please see attached excel document, “What was done,” for details. We had multiple phases of data cleaning. These data represented one of our initial phases.

We corrected all typos/spelling mistakes; however, I note that there are four still present in this version: “GOFFER” (should be GOPHER or GOLFER), “PLIERSS” (should be PLIERS), “HOTDOG” (should be HOT DOG), “PEANUT-BUTTER” (should be PEANUT BUTTER).

In this version, we converted many responses from plural form to singular (see blue columns in attached excel sheet). However, please note that not all plural responses were converted to singular (see orange columns in attached excel sheet). Thus, I think it would benefit to re-run the analysis after cleaning the flagged responses in the orange column and the typos.   
  
Overall, I’m not sure how you’d like to count these changes – by number of words converted or by number of responses converted (remember, some words appear multiple times across conditions). If you count this change by the number of responses changed from plural to singular, then these add up to 244 responses overall (including duplicated word changes). If you count this change by the number of unique words, then this total is much smaller: 43 unique words. Please double check/re-calculate these numbers using the spreadsheet, after you’ve completed data cleaning.

Please note that when administering the WAT, research staff were instructed to cue participants for an alternative response if they gave any of the following types of responses: proper nouns (e.g., “KEVIN” in response to stimulus, HAIRDRESSER), phrases (e.g., “NOT A FINGER” in response to cue, THUMB), non-English word, multi-word responses (e.g., “GENETICALLY ENGINEERED”), and numerical answers (e.g., “TWENTY” in response to the cue, HAIRDRESSER). These data represent the participants’ first response given, regardless of their validity of their response. Thus, we accepted all kinds of responses for the semantic analysis. We have no invalid responses flagged in these data.  
  
In addition, we did not collapse similar responses together for this version. For example, "SWIMMING" and "SWIM" were treated as unique responses as well as "SQUEEZE" and "SQUEEZED.”

**DEMOGRAPHICS FOR WAT DATA**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Visual Artists (VIS)** | **Scientists (SCI)** | **Smart Comparison Group (SCG)** |
| **Age (*M*, *SE*)** | 44.30 (1.38) | 46.72 (1.35) | 41.77 (1.76) |
| **Male *n*** | 18 | 18 | 14 |
| **Female *n*** | 15 | 19 | 14 |
| **Total *N*** | 33 | 37 | 28 |
| **Estimated IQ (*M*, *SE*)** | 111.74 (1.89) | 113.97 (1.68) | 113.25 (1.98) |

The following participants of the entire Big-C study sample did NOT complete the WAT: 2 VIS, 4 SCI, 4 SCG.

The following participants who completed the WAT did NOT complete the WAIS-Vocabulary and WAIS-Matrix Reasoning, so their estimated IQ could not be calculated: 0 VIS, 2 SCI, 1 SCG.