

BEHAVIOR RESEARCH
METHODS

**You don't need t- or F-tests to show that your groups are
matched**

Journal:	<i>Behavior Research Methods</i>
Manuscript ID	Draft
Manuscript Type:	Original Manuscript
Date Submitted by the Author:	n/a
Complete List of Authors:	Politzer-Ahles, Stephen; The Hong Kong Polytechnic University, Chinese and Bilingual Studies Chen, Si; The Hong Kong Polytechnic University, Chinese and Bilingual Studies

SCHOLARONE™
Manuscripts

View Only

You don't need *t*- or *F*-tests to show that your groups are matched

Stephen Politzer-Ahles, Si Chen

*Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University,
Hong Kong*

Address correspondence to:

Stephen Politzer-Ahles

Department of Chinese and Bilingual Studies

The Hong Kong Polytechnic University

Hong Kong

E-mail: stephen.politzerahles@polyu.edu.hk

Phone: +852 27667891

For Review Only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Abstract

Researchers who want to show that two groups are matched on some control variable often attempt to do so by using statistical tests (e.g. a *t*-test or analysis of variance) in which the difference between groups is not statistically significant. This approach is useless for the purpose it is usually intended for (i.e., proving that differences in some outcome variable are attributable to theoretical interesting differences between the groups and not attributable to confounds) because inferential statistics are about population effects whereas the question of interest in this situation is about sample differences. Instead of using irrelevant inferential tests to assume that samples are matched, it is more useful to directly include confounding variables in the statistical model.

Keywords: matching, inferential statistics, between-groups designs

Introduction

When psychological research involves comparisons of outcomes between groups (e.g., between participants or between stimuli), researchers often want to be able to show that differences between those groups are due to the manipulation of interest and not to other confounds. For instance, if a researcher measures language ability in children with and without Autism Spectrum Disorder and wants to argue that the disorder is associated with lower language ability, they generally will want to make sure that any differences they observe in language ability cannot

1
2
3 be attributed to other relevant variables such as age or nonverbal intelligence,
4 which might also differ between the groups. To take another example, imagine an
5 experiment in which participants are shown a real words and made-up words (one
6 at a time, in random order) and instructed to indicate as quickly and accurately as
7 possible if each word is real or made-up. If the researcher wants to demonstrate
8 that the speed of word comprehension (as measured by response speed in the task)
9 is different between nouns and verbs, they will want to make sure any differences
10 in response time between nouns and verbs cannot instead be attributed to other,
11 possibly confounding, factors, like how long the words are.
12
13
14
15
16
17
18
19
20
21
22
23

24 Researchers often do this by producing "matched" samples, e.g., finding children
25 with and without Autism Spectrum disorder such that the two groups have similar
26 age and nonverbal intelligence, or selecting nouns and verbs such that the two sets
27 of words are similar in length. Furthermore, researchers often attempt to
28 demonstrate that these samples are well matched by showing non-significant
29 inferential statistical tests comparing the groups on relevant variables. For example,
30 if we were carrying out the abovementioned word recognition study using 36 nouns
31 and 36 verbs, we might make a statement such as the following:
32
33
34
35
36
37
38
39
40
41
42

43 *The mean length in letters was 8.3 (SD=2.1) for nouns and 7.7 (SD=2.3)*
44 *for verbs; this difference was not significant (Welch $t(68.6)=0.90$,*
45 *$p=.372$).*
46
47
48
49
50

51 and then go on to be confident that any reaction time differences we observe
52 between nouns and verbs cannot be due to the words' length.
53
54
55
56
57
58
59
60

1
2
3 This use of inferential statistics is common in much research in experimental
4 psychology. For instance, we reviewed the latest volume (2017) of *Language,*
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

This use of inferential statistics is common in much research in experimental psychology. For instance, we reviewed the latest volume (2017) of *Language, Cognition and Neuroscience*, a journal in psycholinguistics and neurolinguistics, and found that out of 63 regular articles presenting empirical experiments, 48% had at least one instance of using an inferential statistical test to conclude that samples of materials or participants were matched or were different (Supplementary File 1).

This is likely an underestimate of how often this is done in between-groups designs (since some of the papers that did not include such statistics were reporting experiments with fully within-stimuli and within-participants designs, which would not necessitate matching anyway). While this analysis only involved one journal, the pattern is consistent with our experience with papers many other journals in psycholinguistics, phonetics, and experimental psychology. In short, inferential statistical tests are widely used by researchers wanting to show that their samples are or are not matched on some critical dimension.

Such inferential tests, however, are useless for demonstrating what the researcher is trying to demonstrate. Below we explain why these tests are useless, and offer some suggestions for better approaches to this problem.

Why inferential statistics don't show that samples are matched

These tests provide inferences about population effects, when the researcher is interested in samples. Inferential statistics are meant to evaluate hypotheses

1
2
3 about populations (Hoekstra et al., 2006).¹ For example, the t -test comparing noun
4 and verb samples tests the null hypothesis that the 36 nouns and 36 verbs came
5 from a population in which nouns and verbs have the same length; the p -value
6 of .372 indicates that if they were indeed randomly sampled from such a population,
7 there is a 37.2% chance that the observed length difference between nouns and
8 verbs could have been obtained in a random sample. The properties of the
9 population, however, are irrelevant to the research question at hand.² In this
10 example, the researcher wants to know if the difference in the *sample* is enough to
11 cause meaningful difference in reaction times; a statistical test providing inferences
12 about the population does not answer this question.
13
14
15
16
17
18
19
20
21
22
23
24
25
26

27 This issue alone is enough to reject the use of inferential tests for
28 demonstrating that groups are matched. Nevertheless, there is another problem
29 with this approach as well: the even if a difference between groups is not
30
31
32
33
34

35
36 ¹ Inferential statistics are most often based on p -values, or on confidence intervals (which
37 are directly mathematically related to p -values). A p -value tests the compatibility between a
38 sample statistic and a statistical model (Wasserstein & Lazar, 2016). In practice, this model
39 is usually a hypothetical population in which the relevant statistic is zero.

40 ² A potential exception to this point is that properties of a sample can be relative to a
41 population. For example, imagine that a researcher finds a sample of nouns and a sample of
42 verbs such that the two samples have the exact same average length, but in the language
43 at large verbs tend to be shorter. In this case, the nouns and verbs have the same absolute
44 length but different relative lengths: the verbs are long for verbs and the nouns are short
45 for nouns. It necessarily follows that, if two groups of items or subjects are matched on
46 some absolute measure, they are not matched on the relative measure. There are indeed
47 contexts where the relative measures may be more important (e.g., frequencies of lexical
48 tones relative to a given syllable: Wiener & Ito, 2015). In any case, however, the kinds of
49 statistical tests described here do not solve this problem, and indeed attempts to match
50 samples mask the problem.
51

52 Another case in which inferential tests may be relevant is for judging the
53 effectiveness of a matching procedure, as opposed to the matched-ness of a sample itself.
54 For example, propensity score matching is often used in observational studies to generate
55 matched samples out of a larger sample (Austin, 2011); inferential statistics can be used to
56 test (and hopefully fail to reject) the hypothesis that the procedure itself generates samples
57 that are the same on some characteristic. The issue remains, however, that this would not
58 be an inference about the sample itself.
59
60

1
2
3 statistically significant, it may still be enough to substantially influence the results
4
5 of interest.
6
7

8
9 *Statistical non-significance does not entail practical non-importance.* Most
10 researchers are familiar with the admonition that "statistical significance does not
11 entail practical importance" (Wasserstein & Lazar, 2016). Likewise, statistical non-
12 significance does not entail practical unimportance. Even if the difference in length
13 between a sample of nouns and a sample of verbs is non-significant, it may still be
14 enough to trigger significant differences in some dependent variable that it is
15 associated with, like reaction times. Supplementary File 2 includes several
16 simulated datasets which all have this property: each dataset has two samples
17 (verbs and nouns) with a non-significant difference in length and a significant or
18 marginal difference in reaction speed. The common approach described above—
19 showing that the difference in length is non-significant and then attributing the
20 difference in reaction time to the manipulation of interest—would lead the
21 researcher to conclude that the difference in reaction time must be due to
22 something other than length (e.g., to an underlying difference between verbs and
23 nouns). In fact, the data were simulated such that the differences in reaction time
24 would be attributable to length. Specifically, data with an $r=.35$ correlation between
25 length and reaction time were simulated, and then grouped arbitrarily into two
26 samples such that nouns had a slightly, but not significantly, higher mean length
27 and a significantly higher mean reaction latency. Regressing reaction time on both
28 part of speech (noun vs. verb) and length, as shown in the right-hand side of Figure
29 1, reveals that the reaction time difference between nouns and verbs is not
30 significant once length is taken into account. In this case, rather than concluding
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

that nouns and verbs are intrinsically associated with different reaction speeds, the researcher might instead conclude that apparent differences between nouns and verbs in terms of reaction speed might just be an epiphenomenon of differences in length (i.e., that nouns and verbs may both fall along the same length – reaction time regression line).

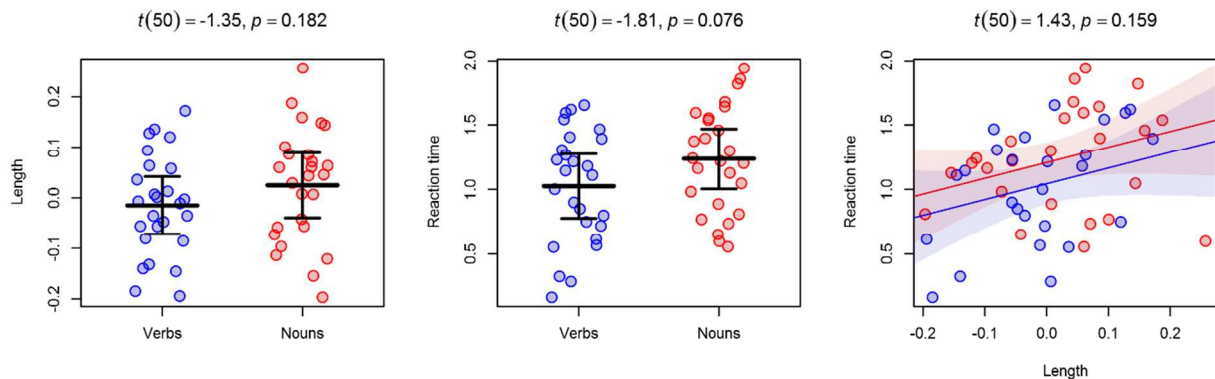


Figure 1. Simulated noun/verb reaction time data. Each point represents one word (red for nouns, blue for verbs). At the top of each panel is shown a t-test comparing the y-axis variable between nouns and verbs (i.e., the t-test on the left panel tests whether nouns and verbs have the same length, and the tests on the center and right panels test whether nouns and verbs have the same reaction time). The left panel shows that the two samples do not significantly differ in length; the long black lines show the mean, and the error bars show difference-adjusted 95% confidence intervals. The center panel shows that the two samples do marginally differ in reaction time. The rightmost panel shows a regression of reaction time on length and part of speech; shaded regions represent the confidence interval of each line. In this panel, the difference between nouns and verbs is no longer significant. While the first two panels suggest that the two samples differ in reaction time, the

1
2
3 *regression suggests that the apparent differences between the samples in reaction*
4 *time are probably due to the confounding difference between samples in length.*
5
6
7
8
9

10 11 12 13 14 **Better ways to deal with control variables** 15

16
17 If providing a non-significant *t*-test or analysis of variance does not help
18 prove that the experiment results are not due to confounds, what else can a
19 researcher do? Surely researchers cannot simply ignore confounds entirely.
20 Nonetheless, a non-significant inferential test does nothing to prove that the
21 confound has been taken care of, and researchers should not mistake such a test as
22 a license to stop worrying about potential confound. Simply describing the data
23 distribution for each sample provides more relevant information than the inferential
24 statistics without descriptions do. E.g., in the example above, if the *t*-test were
25 removed and the means and standard deviation left, the reader would not lose
26 anything that helps them understand the results. Thus, simply providing the
27 descriptive statistics (or, better yet, the full list of stimuli/participants and relevant
28 variables for each; Morey et al., 2016; Vazire, 2017) and letting the reader judge
29 for themselves whether or not there is a problem, is more useful than simply
30 providing an inferential test. Furthermore, it avoids giving anybody the false
31 impression that a confound has been eliminated.
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

51
52 If descriptive statistics are not enough to satisfy a researcher or their readers
53 that the results of a study are not due to some confounding variable, what more
54 can be done? In such a case, instead of resorting to an irrelevant inferential
55
56
57
58
59
60

1
2
3 statistic, researchers can instead include the confounding variables in their
4
5 statistical model, or researchers with directional hypotheses can select samples
6
7 where confounds cannot account for results of interest.
8
9

10
11 *Including control variables in the statistical model.* The best option is to
12
13 explicitly account for relevant variables in the statistical model. A regression model
14
15 can take into account many variables at once and try to identify whether one
16
17 variable of interest has a unique association with the outcome variable when other
18
19 variables are accounted for (Keith, 2006; Westfall & Yarkoni, 2016). For example,
20
21 the researcher in the word recognition experiment could construct one model
22
23 predicting reaction times as a function of length and word category (as well as
24
25 relevant random effects; Baayen et al., 2008, Judd & Westfall, 2012), and another
26
27 model predicting reaction times as a function of length only. The researcher would
28
29 then compare the two models' goodness of fit to test whether the inclusion of word
30
31 category significantly improves how well the model fits the data (Keith, 2006). If it
32
33 does, the researcher can infer that noun/verb status has a significant association
34
35 with reaction times over and above the confounded effect of length. On the other
36
37 hand, if the inclusion of this factor does not significantly improve the model fit, the
38
39 researcher does not have evidence to make such a conclusion; this is the case for
40
41 all datasets in Supplementary File 2. Indeed, taking multiple and potentially
42
43 confounded variables into account at once is one of the main purposes of regression
44
45 models (Christenfield et al., 2004; Keith, 2006). An additional benefit of this
46
47 practice is that it obviates the need to form matched samples, which can be quite
48
49 difficult when the number of relevant control variables is high (Balota et al., 2004;
50
51 Pourhoseingholi et al., 2012).
52
53
54
55
56
57
58
59
60

1
2
3 This solution is not perfect, because "regressing out" some control variable
4 like is not the same as regressing out the actual confounding *construct* that is really
5 influencing reaction times (Westfall & Yarkoni, 2016).³ For example, length in
6 letters is only a rough measure of how "big" a word actually is on the page or
7 computer screen; if words are written in a font that is not monospace, words with
8 the same length in letters might take up more space when typed out on the page.
9 While this example is, by design, very simple, in actual research many constructs
10 we are interested in, like personality, working memory capacity, language
11 proficiency, how frequently a word is used, etc., are unobservable and can only be
12 estimated using rough measures. While this poses a challenge for researchers
13 trying to use regression models to control for confounding variables (Westfall &
14 Yarkoni, 2016), the exact same shortcoming applies to studies trying to create
15 samples of stimuli that are matched on length or some other confounding variable:
16 the samples can only be matched on a rough measure, not on the true construct,
17 which is un-observable.
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37

38 Another limitation of this approach is that some models may not be able to
39 account well for the effect of the confounding variable. For example, if our sample
40 of words has very little variation in length, it may be difficult to model this variable
41 well; likewise, it may be difficult to fit a model to the length effect if a researcher
42 originally dichotomized the variable by choosing samples of long and short words,
43 thus ending up with a sample in which length is bimodally distributed. Furthermore,
44 some such models may be difficult to implement. For example, if a manipulation is
45 both between participants and between stimuli, and both of these random factors
46
47
48
49
50
51
52
53
54
55

56
57
58
59
60
3

1
2
3 has potentially confounding variables, then a mixed model (Baayen et al., 2008)
4
5 would be needed to account for all of these variables together, and this may be
6
7 difficult to apply to situations which require some sort of aggregation over
8
9 participants or stimuli first (such as cluster-based permutation tests used on event-
10
11 related potential data; Maris & Oostenveld, 2007). In cases like this, putting the
12
13 confounding variable into the model may not solve the problems, and the
14
15 researcher might have no choice but to only offer descriptive statistics, as
16
17 mentioned above. This would still be better than offering a non-significant
18
19 inferential statistic that gives the researcher false confidence that the groups are
20
21 well matched.
22
23
24
25
26

27 *Taking advantage of directional hypotheses.* Another option researchers can
28
29 use if they are unable to use a regression model is to simply make sure that any
30
31 difference between groups is not confounded with their predicted result. In the
32
33 word recognition example, imagine that the researcher knows that less-frequently-
34
35 used words are usually recognized more slowly, *and* the researcher's hypothesis is
36
37 that nouns will be recognized more slowly than verbs. If the researcher intentionally
38
39 samples stimuli such that the nouns are more frequently used than verbs,⁴ *and*
40
41 observes slower reaction times for nouns in the experiment, then the reaction time
42
43 difference cannot be attributed to length. This approach, however, requires having
44
45 a directional research hypothesis, and requires great confidence in the direction of
46
47 the possibly confounding effect (e.g., that common words are responded to more
48
49
50
51

52 ⁴ Or used with the exact same frequency as verbs, but this is usually impossible. For control
53
54 variables that are continuous, no two observations can ever be exactly the same. For
55
56 instance, if two people were born at the "same" time down to the second, they still do not
57
58 have the exact same age; someone using a more fine-grained clock (e.g. with nanosecond-
59
60 level resolution) would eventually be able to detect a small difference in time of birth when
measuring with sufficient detail.

1
2
3 quickly, not more slowly). Therefore, it is not applicable in situations where either
4
5 the outcome of interest in the experiment, or the impact of the confounding
6
7 variable, might be expected to go in either direction.
8
9

10 11 12 13 14 **Conclusion**

15
16
17 While researchers are right to worry about confounding variables, the
18
19 common approach of using non-significant t - or F -tests to try to show that groups
20
21 are "matched" on these variables does not resolve this problem, as it provides no
22
23 information germane to the question of how these confounds are influencing the
24
25 effects of interest. Instead of trying to use such tests to claim that groups are
26
27 matched, it is more productive to directly include these confounding variables in the
28
29 statistical model.
30
31
32
33
34
35

36 37 **References**

38
39 Austin, P. (2011). An introduction to propensity score methods for reducing the
40
41 effects of confounding in observational studies. *Multivariate Behavioral Research*,
42
43 46, 399-424.
44
45

46
47 Baayen, H., Davidson, D., & Bates, D. (2008). Mixed-effects modeling with crossed
48
49 random effects for subjects and items. *Journal of Memory and Language*, 59, 390-
50
51 412.
52
53
54
55
56
57
58
59
60

1
2
3 Balota, D., Cortese, M., Sergent-Marshall, S., Spieler, D., & Yap, M. (2004). Visual
4 word recognition for single-syllable words. *Journal of Experimental Psychology:*
5
6 *General*, 133, 283-316.
7

8
9
10 Christenfeld, N., Sloan, R., Carroll, D., & Greenland, S. (2004). Risk factors,
11 confounding, and the illusion of statistical control. *Psychosomatic Medicine*, 66,
12
13 868-875.
14
15

16
17
18 Hoekstra, R., Kiers, H., Johnson, A., Groenier, M. (2006) Problems when
19 interpreting research results using only p-value and sample size. In: *ICOTS-7*
20
21 *International Conference on Teaching Statistics: Working Cooperatively in Statistics*
22
23 *Education*.
24
25

26
27
28 Judd, C., Westfall, J., & Kenny, D. (2012). Treating stimuli as a random factor in
29 social psychology: a new and comprehensive solution to a pervasive but largely
30
31 ignored problem. *Journal of Personality and Social Psychology*, 103, 54-69.
32
33

34
35
36 Keith, T. (2006). *Multiple regression and beyond*. Boston: Pearson.
37

38
39 Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and
40
41 MEG-data. *Journal of Neuroscience Methods*, 164, 177-190.
42
43

44
45 Morey, R., Chambers, C., Etchells, P, Harris, C., Hoekstra, R., Lakens, D., ..., &
46
47 Zwaan, R. (2016). The Peer Reviewers' Openness Initiative: incentivising open
48
49 research practices through peer review. *Royal Society Open Science*, 3, 150547.
50

51
52 Pourhoseingholi, M., Baghestani, A., & Vahedi, M. (2012). How to control
53
54 confounding effects by statistical analysis. *Gastroenterology and Hepatology: From*
55
56 *Bed to Bench*, 5, 79-83.
57
58
59
60

1
2
3 Vazire, S. (2017). Quality uncertainty erodes trust in science. *Collabra*, 3, 1.
4

5
6 Wasserstein, R., & Lazar, N. (2016). The ASA's statement on p-values: context,
7
8 process, and purpose. *The American Statistician*, 70, 129-133.
9

10
11 Westfall, J., & Yarkoni, T. (2016). Statistically controlling for confounding constructs
12
13 is harder than you think. *PLoS ONE*, 11, e0152719.
14

15
16 Wiener, S., & Ito, K. (2015). Do syllable-specific tonal probabilities guide lexical
17
18 access? Evidence from Mandarin, Shanghai and Cantonese speakers. *Language,*
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
60
Cognition, and Neuroscience, 30, 1048-1060.

For Review Only

	Paper	Journal	JournalVolume	Year
1				
2				
3	Brown-Schmidt and Toscano	LCN	LCN32_10	2017
4	Henry et al.	LCN	LCN32_10	2017
5	Middleton et al.	LCN	LCN32_10	2017
6				
7	Martín-Loeches et al.	LCN	LCN32_10	2017
8	Dufour et al.	LCN	LCN32_10	2017
9	Kapnoula et al.	LCN	LCN32_10	2017
10	Wong et al.	LCN	LCN32_10	2017
11	Arantzeta et al.	LCN	LCN32_10	2017
12	Grieco-Calub	LCN	LCN32_10	2017
13	Caffarra et al.	LCN	LCN32_9	2017
14	Scaltritti et al.	LCN	LCN32_9	2017
15	Shahin et al.	LCN	LCN32_9	2017
16	Croot et al.	LCN	LCN32_9	2017
17	Mitterer and Reinisch	LCN	LCN32_9	2017
18				
19	Yue et al.	LCN	LCN32_9	2017
20	Yu et al.	LCN	LCN32_9	2017
21				
22	Lawyer and Corina	LCN	LCN32_9	2017
23	Frank and Willems	LCN	LCN32_9	2017
24	Bellanger et al.	LCN	LCN32_9	2017
25	Ito et al.	LCN	LCN32_8	2017
26	van Rijswijk et al.	LCN	LCN32_8	2017
27	Calloway and Perfetti	LCN	LCN32_8	2017
28	Turnbull et al.	LCN	LCN32_8	2017
29	Fieder et al.	LCN	LCN32_8	2017
30	Gu et al.	LCN	LCN32_8	2017
31	Liu et al.	LCN	LCN32_6	2017
32				
33	Bürki	LCN	LCN32_6	2017
34	Gould et al.	LCN	LCN32_6	2017
35	Vivas et al.	LCN	LCN32_6	2017
36				
37	Bühler et al.	LCN	LCN32_6	2017
38	Zellou et al.	LCN	LCN32_6	2017
39	Akhavan et al.	LCN	LCN32_6	2017
40	Ghitza	LCN	LCN32_5	2017
41	Rommers et al.	LCN	LCN32_5	2017
42	Lam et al.	LCN	LCN32_5	2017
43	Lewis et al.	LCN	LCN32_5	2017
44				
45	Tsang et al.	LCN	LCN32_5	2017
46	Fang et al.	LCN	LCN32_5	2017
47	Monaco et al.	LCN	LCN32_5	2017
48	Babineau et al.	LCN	LCN32_4	2017
49	Reifegerste et al.	LCN	LCN32_4	2017
50	Declerck and Philipp	LCN	LCN32_4	2017
51	Olson	LCN	LCN32_4	2017
52	Smolka and Libben	LCN	LCN32_4	2017
53	Riesenhuber and Glezer	LCN	LCN32_3	2017
54	Gubian et al.	LCN	LCN32_3	2017
55	Vankov and Bowers	LCN	LCN32_3	2017
56				
57				
58				
59				
60				

1				
2	Brehm and Bock	LCN	LCN32_2	2017
3	Schiller et al.	LCN	LCN32_2	2017
4	Weisberg et al.	LCN	LCN32_2	2017
5	Ivanova et al.	LCN	LCN32_2	2017
6	Borovsky	LCN	LCN32_2	2017
7	Pycha	LCN	LCN32_2	2017
8	Zhuang and Devereux	LCN	LCN32_2	2017
9	Bishop	LCN	LCN32_2	2017
10	Miwa et al.	LCN	LCN32_1	2017
11	Achim et al.	LCN	LCN32_1	2017
12	Chen and Yeh	LCN	LCN32_1	2017
13	Rose and Rahman	LCN	LCN32_1	2017
14	Tsang et al.	LCN	LCN32_1	2017
15	Shantz and Tanner	LCN	LCN32_1	2017
16	Wu et al.	LCN	LCN32_1	2017
17	Hwang	LCN	LCN32_1	2017
18				
19				
20				
21				
22				
23				
24				
25				
26				
27				
28				
29				
30				
31				
32				
33				
34				
35				
36				
37				
38				
39				
40				
41				
42				
43				
44				
45				
46				
47				
48				
49				
50				
51				
52				
53				
54				
55				
56				
57				
58				
59				
60				

1 has inferential statistic for making conclusions about samples?
2
3

- 4 0
- 5 1
- 6 0
- 7 0
- 8 0
- 9 1
- 10 1
- 11 0
- 12 0
- 13 1
- 14 1
- 15 0
- 16 0
- 17 0
- 18 0
- 19 0
- 20 0
- 21 1
- 22 1
- 23 0
- 24 1
- 25 1
- 26 1
- 27 1
- 28 1
- 29 1
- 30 0
- 31 0
- 32 1
- 33 1
- 34 1
- 35 1
- 36 1
- 37 1
- 38 1
- 39 0
- 40 0
- 41 0
- 42 0
- 43 0
- 44 0
- 45 0
- 46 1
- 47 1
- 48 0
- 49 1
- 50 0
- 51 1
- 52 1
- 53 1
- 54 1
- 55 0
- 56 0
- 57 0
- 58 0
- 59 0
- 60

For Review Only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
1
0
0
0
0
0
1
1
0
1
0
0
0
1
1

For Review Only

	Y	x	cond
1			
2			
3	0.504044	0.115988	A
4	0.597747	0.183254	A
5	0.605813	-0.03392	A
6	0.721443	-0.1191	A
7	0.738369	-0.14843	B
8	0.749678	0.225042	A
9	0.770495	-0.03131	B
10	0.781449	-0.02786	A
11	0.795879	-0.13236	B
12	0.801702	0.027689	A
13	0.802044	0.066873	B
14	0.905733	-0.07725	A
15	0.926484	0.029461	B
16	0.929668	0.10807	A
17	0.957128	-0.07128	B
18	0.963321	-0.07461	A
19	0.99234	-0.09293	B
20	0.999539	0.204705	A
21	1.018425	-0.01128	B
22	1.022739	-0.03652	A
23	1.023101	-0.01877	B
24	1.026082	-0.11643	A
25	1.039974	-0.15389	B
26	1.070918	-0.19812	A
27	1.090003	-0.00791	B
28	1.13825	0.071231	A
29	1.143477	-0.09521	B
30	1.185487	0.006076	A
31	1.186985	-0.00045	B
32	1.256148	-0.0502	A
33	1.258514	-0.02055	B
34	1.28407	0.003	A
35	1.286365	0.090335	B
36	1.30271	0.011283	A
37	1.309562	0.096223	B
38	1.330447	-0.14638	A
39	1.417858	0.017078	B
40	1.472239	-0.26579	A
41	1.525241	-0.02916	B
42	1.574016	-0.14771	A
43	1.651307	0.160608	B
44	1.669936	-0.08774	A
45	1.727708	0.033348	B
46	1.76846	0.078238	A
47	1.855449	-0.0568	B
48	1.899281	-0.05975	A
49			
50			
51			
52			
53			
54			
55			
56			
57			
58			
59			
60			

For Review Only

1		
2	1.933937	-0.00858 B
3	1.945686	0.087497 A
4	1.977582	0.04809 B
5	2.036047	0.159971 B
6	2.153764	0.076605 B
7	2.170337	0.075684 B
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		
25		
26		
27		
28		
29		
30		
31		
32		
33		
34		
35		
36		
37		
38		
39		
40		
41		
42		
43		
44		
45		
46		
47		
48		
49		
50		
51		
52		
53		
54		
55		
56		
57		
58		
59		
60		

For Review Only

	Y	x	cond
1			
2			
3	0.423114	0.032769	A
4	0.540773	-0.1018	A
5	0.603295	0.101789	A
6			
7	0.66598	0.111585	A
8	0.669595	-0.16077	B
9	0.673531	0.031569	A
10	0.698546	-0.12798	B
11	0.778257	-0.13825	A
12	0.790263	0.176789	B
13			
14	0.79949	0.152493	A
15	0.814503	0.007965	B
16	0.829434	0.007874	A
17	0.838239	0.121257	B
18	0.841395	-0.1127	A
19	0.866606	0.006526	B
20	0.938719	-0.12751	A
21	0.953075	0.037643	B
22	0.960099	-0.17012	A
23	0.985214	0.023912	B
24			
25	1.02132	0.026639	A
26	1.048531	-0.05399	B
27	1.081667	0.186066	A
28	1.101311	-0.13764	B
29	1.112555	-0.02012	A
30	1.148615	0.080326	B
31	1.153951	-0.12871	A
32	1.171355	0.08466	B
33	1.197304	-0.00218	A
34	1.239122	-0.14545	B
35	1.241997	0.154367	A
36	1.248006	-0.00173	B
37	1.253114	-0.08527	A
38	1.264995	-0.01879	B
39			
40	1.27206	-0.0637	A
41	1.303732	0.258468	B
42	1.314489	-0.11428	A
43	1.351257	0.137476	B
44	1.369353	0.112902	A
45	1.376125	-0.09039	B
46	1.395906	-0.0368	A
47	1.416968	0.126259	B
48	1.435761	0.068511	A
49	1.460717	-0.02423	B
50	1.483648	-0.02606	A
51	1.500215	0.182193	B
52	1.503537	0.02982	A
53			
54			
55			
56			
57			
58			
59			
60			

For Review Only

1			
2	1.58792	0.031239	B
3	1.59541	-0.15495	A
4	1.601028	0.027073	B
5	1.618038	0.284162	B
6	1.706993	-0.01348	B
7	1.732961	0.042319	B
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
32			
33			
34			
35			
36			
37			
38			
39			
40			
41			
42			
43			
44			
45			
46			
47			
48			
49			
50			
51			
52			
53			
54			
55			
56			
57			
58			
59			
60			

For Review Only

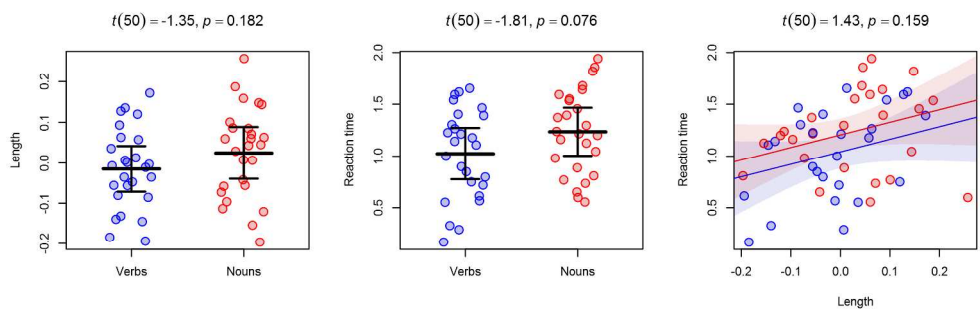
	Y	x	cond
1			
2			
3	0.164284	-0.18472	A
4	0.286197	0.006731	A
5	0.325619	-0.13995	A
6	0.552318	0.03579	A
7	0.554854	0.060489	B
8	0.566297	-0.01083	A
9	0.598027	0.257287	B
10	0.61398	-0.19421	A
11	0.65203	-0.04182	B
12	0.719211	-0.00285	A
13	0.736169	0.070882	B
14	0.750403	0.120045	A
15	0.769202	0.100609	B
16	0.798556	-0.03513	A
17	0.809214	-0.19671	B
18	0.849928	-0.04712	A
19	0.887225	0.007683	B
20	0.899017	-0.05598	A
21	0.981483	-0.07275	B
22	1.000652	-0.00683	A
23	1.04609	0.143729	B
24	1.108824	-0.14558	A
25	1.125854	-0.15454	B
26	1.143608	-0.132	A
27	1.161747	-0.09627	B
28	1.178305	0.057562	A
29	1.200934	-0.12091	B
30	1.213363	0.001411	A
31	1.217103	-0.05572	B
32	1.227018	-0.05559	A
33	1.238522	-0.11366	B
34	1.264714	0.063227	A
35	1.296649	0.006849	B
36	1.303731	-0.08056	A
37	1.37233	-0.0578	B
38	1.391424	0.172062	A
39	1.394786	0.086299	B
40	1.404171	-0.03514	A
41	1.45626	0.158858	B
42	1.465307	-0.08542	A
43	1.536146	0.187418	B
44	1.541001	0.093317	A
45	1.551898	0.029006	B
46	1.593013	0.127521	A
47	1.593303	0.059733	B
48	1.618619	0.13547	A
49			
50			
51			
52			
53			
54			
55			
56			
57			
58			
59			
60			

For Review Only

1		
2	1.641425	0.084915 B
3	1.653108	0.012852 A
4	1.67784	0.043443 B
5	1.818058	0.147876 B
6	1.856941	0.045413 B
7	1.941792	0.062975 B
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		
25		
26		
27		
28		
29		
30		
31		
32		
33		
34		
35		
36		
37		
38		
39		
40		
41		
42		
43		
44		
45		
46		
47		
48		
49		
50		
51		
52		
53		
54		
55		
56		
57		
58		
59		
60		

For Review Only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



Or Review Only