

A Replication of Rothman (2011): L3 syntactic transfer selectivity and typological
determinacy: The typological primacy model

Kyle Parrish¹

¹ Rutgers University

Author Note

Add complete departmental affiliations for each author here. Each new line herein
must be indented, like this line.

Enter author note here.

Correspondence concerning this article should be addressed to Kyle Parrish, 15
Seminary Place, New Brunswick, NJ. E-mail: kyle.parrish@rutgers.edu

Abstract

One or two sentences providing a **basic introduction** to the field, comprehensible to a scientist in any discipline.

Two to three sentences of **more detailed background**, comprehensible to scientists in related disciplines.

One sentence clearly stating the **general problem** being addressed by this particular study.

One sentence summarizing the main result (with the words “**here we show**” or their equivalent).

Two or three sentences explaining what the **main result** reveals in direct comparison to what was thought to be the case previously, or how the main result adds to previous knowledge.

One or two sentences to put the results into a more **general context**.

Two or three sentences to provide a **broader perspective**, readily comprehensible to a scientist in any discipline.

Keywords: keywords

Word count: X

A Replication of Rothman (2011): L3 syntactic transfer selectivity and typological
determinacy: The typological primacy model

The distinction between third language (L3) and second language (L2) learning in adulthood has received a significant increase in research over the last 15 years. One key difference between L3 and L2 learning is that third language learners can be influenced by two languages during the process of acquisition, rather than just their native language as is the case in L2 learning. However, there is not a consensus how exactly and if L3 learners are impacted by their two known languages. Guided by empirical research, several models have been proposed to predict how the L1 and L2 will impact L3 learning and acquisition. One of the first models suggested that L3 learners could draw on either their first or second language, and that this influence would be helpful (The Cumulative Enhancement Model; **(flynn?)**). Others predict that the L2 has a privileged status and inhibits access to the L1 (the L2 Status Factor; Bardel & Falk), or, oppositely, that the native language has a special status (Hermas, 2010). Finally, another set of models suggest that structural similarity between the learned language and background language(s) predicts influence. The Typological Primacy Model (TPM) predicts that wholesale (one entire language) transfer occurs to the L3 at some point during the “initial stages” of L3 development. The language which transfers, either the L1 or the L2, is determined by the individual on the basis of input, but is generally predicted to be the language that is structurally closer to the L3 (e.g., Spanish, not English, would transfer to L3 Brazilian Portuguese). The Linguistic Proximity Model (LPM) also posits that structural similarity plays a key role predicting how the L1 or L2 influence the L3, but suggests that this happens in a property-by-property basis, rather than all at once. These various predictions have each received some support in the literature. In a recent systematic review, **(rothman?)**/**(puig?)** reviewed 92 total L3 studies, aiming to aggregate the evidence for each model. The results showed that, while each model had some empirical support, most studies could be explained by Typological transfer as predicted by the TPM. Many studies also were potentially explained, however, by L2 status effects. It is

important to note that many of these studies were coded as providing evidence simultaneously for two views. That is, a single study which reported L2 influence on the L3 was coded as both L2 and typological influence, since they could not demonstrate that the influence in the L3 was due to L2 status and not structural similarity.

In order to avoid this confound, it has been suggested that L3 studies utilize a so-called mirror-image group design. The mirror image design, proposed in ((foote?), Rothman 2010), recruits two groups that speak the same L3, and same background languages, but their order of acquisition is reversed for the L1 and L2 (e.g. L1 Spanish-L2 English-L3 BP, L1 English-L2 Spanish-L3 BP). By comparing the L3 of these groups, distinct performance on a given experimental task would provide evidence that order of acquisition is an important predictor in predicting how background languages influence a third language. On the other hand, if the groups perform similarly and L2 proficiency is taken into account, it would be evidence that having the same background languages, regardless of their order of acquisition, would predict L3 transfer.

One of the first studies to utilize this mirror image design was Rothman (2011). In the study, two L3 groups, (L1 Italian-L2 English-L3 Spanish and L1 English-L2 Spanish-L3 Brazilian Portuguese) along with a Spanish native and Brazilian Portuguese native group, completed two tasks which measured their judgment of determiner phrase noun-adjective combinations. In Spanish, Italian, and Portuguese, adjectives can appear in pre-nominal or post-nominal position and constitute a change in meaning. Figure *mark* from Rothman (2011) provides an example; post-nominal adjectives induce a set-denoting interpretation, while pre-nominal adjectives elicit a kind-denoting interpretation. English, on the other hand, does not distinguish kind and set denotation syntactically, as the adjective appears pre-nominally in both bases. The results of both tasks revealed that all four groups similarly judged sentences given to them in either their L3 or L1 (for the native groups). This was taken as evidence that, since one group's had this feature and the other group's L2 had the

L3 features, that the structural similarity of these languages, rather than the order of acquisition, predicted L3 transfer.

The present study

The present study is a conceptual replication of Rothman (2011), in which the original methods and analysis were intended to be replicated in a similar population of L3 learners. This replication re-created similar stimuli using the available examples from the original study, and other similar studies that used the same type of tasks (cites), and measured L2 and L3 proficiency using the LexTALE, rather than a cloze and grammar test.

The rationale for this replication is its importance and impact on the field of L3 acquisition over the last decade and would have implications for both supporters of the TPM and other models. In the event that a difference is found between L3 groups, the basis for the model itself could be called into question in favor of other models. This finding would call into question whether the TPM can be considered the dominant or most accurate model in L3 acquisition, given that the recent systematic review (Rothman et al., 2019) concluded that typological transfer accounted for most outcomes in L3 morphosyntax, rather than order of acquisition. However, many of these studies' designs and statistical analyses only allow for binary interpretation of the results. That is, the design of these studies is such that either the L3 is impacted by the L1 or the L2, rather than allowing for the combination of both languages to impact the L3. This strategy is present in many L3 studies and is arguably influenced by Rothman (2011). A failed replication of this study would not only call into question its own findings, but findings in many other studies which adopted similar methodology and now serve as the empirical basis to the TPM. On the other hand, a successful replication of Rothman (2011) would suggest that full influence is possible and would strengthen the basis of the TPM and lessen counter arguments that include potential sampling issues, while also quantifying the (un)certainty related to this similar performance between groups.

Methods

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study.

Participants

A total of 202 participants took part in this experiment, consisting of 4 total groups. The first two groups were L3 speakers of Portuguese who spoke L1 English-L2 Spanish ($n = 12$; henceforth the L1 English group) and L1 Spanish-L2 English ($n = 60$; henceforth the L1 English group). Additionally, a group of Portuguese native speakers ($n = 61$) and a group of Spanish native speakers (69) also participated.

Materials

Two tasks were given to participants

Semantic interpretation task. This task was designed to evaluate how the adjective's position in a DP (pre or post nominal) impacted meaning in both the L2 and L3. The task had both a Spanish and Brazilian Portuguese version, in which 5 target items were tested in each of a pre-nominal and post-nominal condition. Equal numbers of fillers were used which tested other properties (e.g., anaphora resolution) used in subsequent studies. Below is an example taken of the Semantic interpretation task in Spanish taken directly from Rothman (2011).

Prompt “Los maridos honestos se merecen el respeto de sus mujeres.”

Answer choices

a.) *De todos los maridos que hay solo algunos, los que son honestos merecen el respeto de sus esposas.*

b.) Todo marido se merece el respeto de su esposa porque todo marido es, por ser

marido, honesto.

Prompt “Los valientes Incas tenían mucho éxito.”

Answer choices

a.) Entre los incas había los valientes y los no valientes, así que todo inca que era valiente también tenía éxito

b.) *Ser Inca equivale a ser valiente, así que todo Inca tenía éxito.*

Context-based Collocation Task. A second task elicited production of adjectival DPs by way of context. In the task, participants read a short story and had to fill in a blank at the end of the story with either a pre or post nominal adjective. Below is an example of the Context-based Collocation Task in Spanish taken directly from Rothman (2011).

Example

Mi esposa se llama Magda. Ella es una persona muy amable y cariñosa. Aunque solo tenemos 22 años, hace mucho tiempo que somos amigas. Magda es una vieja amiga _____ (viejo).

‘My best friend is named Magda. She is a very nice and affectionate person. Even though we are only 22-years-old, we have been friends for a long time. Magda is an old friend.’

Procedure

Statistical analysis

Two sets of analyses were carried out, a replication of the analysis of the original study and a second analysis using a Bayesian approach. The original study used two-way ANOVAs for each of the two tasks (the Semantic interpretation and context-based collocation tasks).

In each model, the number of correct responses was analyzed as a function of item type (preposed or postposed), group and their interaction.

A. Why Bayesian. If the audience requires it, explain what benefits will be gleaned by a Bayesian analysis (as opposed to a frequentist analysis). B. Goals of analysis. Explain the goals of the analysis. This prepares the audience for the type of models to expect and how the results will be described.

The goal of the Bayesian analysis was to determine whether there were differences between the probability of a correct response by each group while taking preposed and postposed positions into account. One advantage of a Bayesian approach is the ability to provide evidence for practical equivalence or estimate differences.

For the Bayesian analysis, two logistic multilevel regression models analyzed the probability of answering a given trial correctly as a function of group, item type, and their interaction for each task. *Random intercepts for item and participant were also included.* The models was run using with 4000 iterations of Hamiltonian Monte-Carlo sampling (1000 warm up), across 4 chains and 8 processing cores.

Results

1 shows the average correct responses (out of 5) for the Context-based collocation task by all 4 groups.

2 shows the average correct responses (out of 5) for the Context-based collocation task by all 4 groups.

Discussion

Overall, the present study failed to replicate the null results reported in Rothman (2011). In particular, the L3 groups performed differently in the present study on both the

semantic interpretation and the collocation tasks. Interestingly, the Spanish bilingual group answered correctly less often than the Spanish native group in both conditions in the semantic interpretation task and in the post-nominal condition for the context-based collocation task. One possible explanation for this finding is that their L2, English, influenced their judgments of meaning in their L3. This result would best be explained by the L2 status factor, which predicts that the L2 influences the L3 while inhibiting access to the L1.

Just as in the original study, these results do not necessarily demonstrate that transfer is not property by property, since only a single property was tested.

These results could not be explained by the TPM, differently from the original study. A few key differences seem readily apparent between this data and Rothman (2011). First, the mean number of correct answers in Rothman (2011) was over 4 out of 5, which suggests that it is possible that a ceiling effect could explain the similarity between groups. In other words, the stimuli used in the study might have been easily interpreted by the participants and been a result of their acquisition of the property in BP, rather than influence from Spanish. Another possibility is sampling error - the original study had 60 total participants across 4 groups. While the statistical power and risk for sampling error cannot be calculated using the data reported in the study, it is arguable that this sample is not representative of the population it is intended to represent. The present study, on the other hand, carried out a power analysis and found that **power analysis**.

These results call into question whether, in addition to this study, other studies in L3A similar issues relating to statistical power might be at play. In particular, low sample size (under-powered) studies have a high risk of producing a false negative finding. Figure 3 demonstrates this issue by using simulated data. In particular, small, medium and large effects are simulated. The figure shows that, unless the effect is very large, the probability of a false negative can be higher than any significant result being found in the data.

201 Additionally, the so-called “winners curse” (cite) is an issue with low sample studies.
202 That is, even when an effect is found in an underpowered study, it is very likely
203 over-estimated. Figure 4 shows this, again using simulated data. Overall, 2000 simulations
204 were carried out.

References

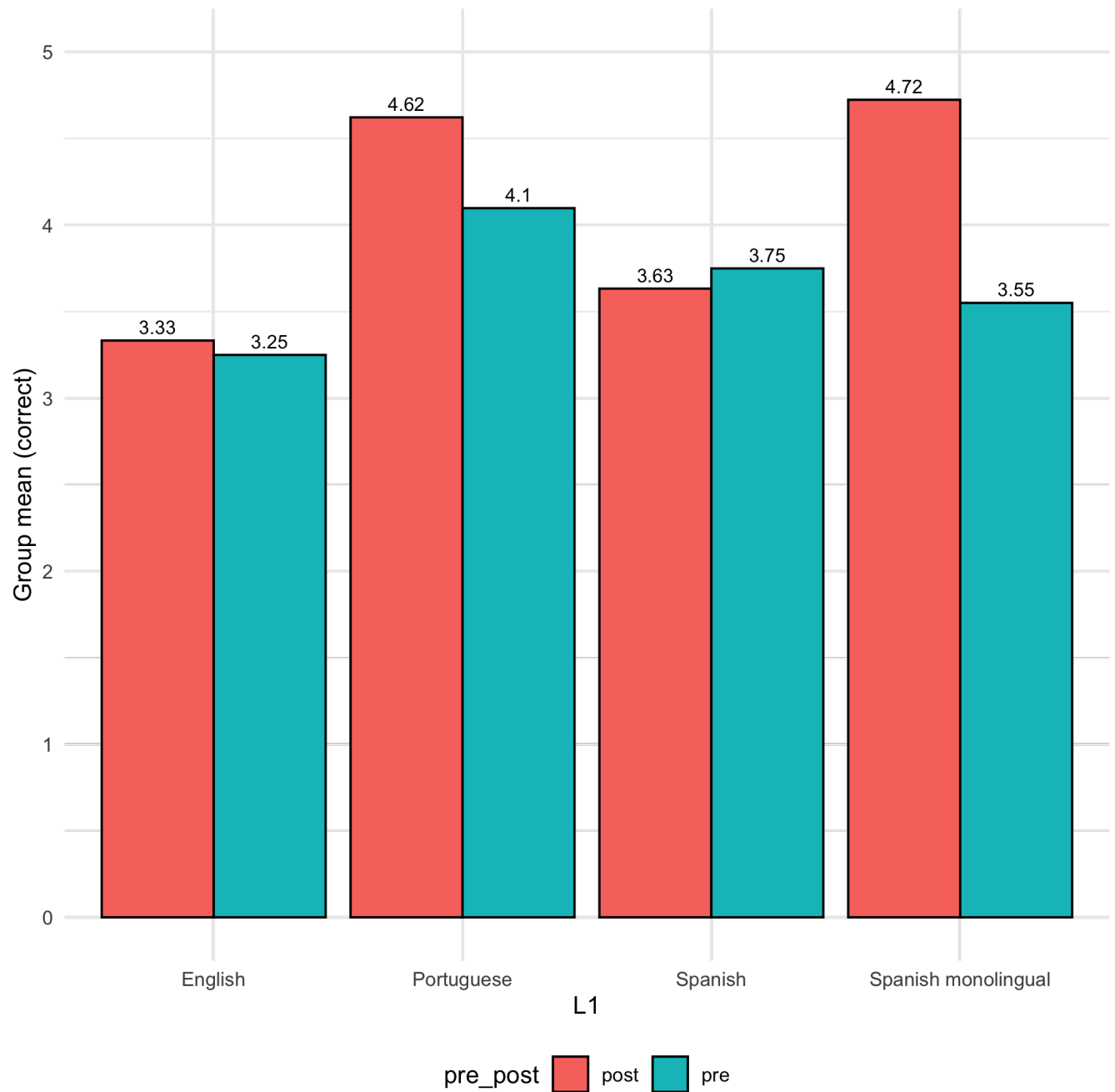


Figure 1. Context-based Collocation Task

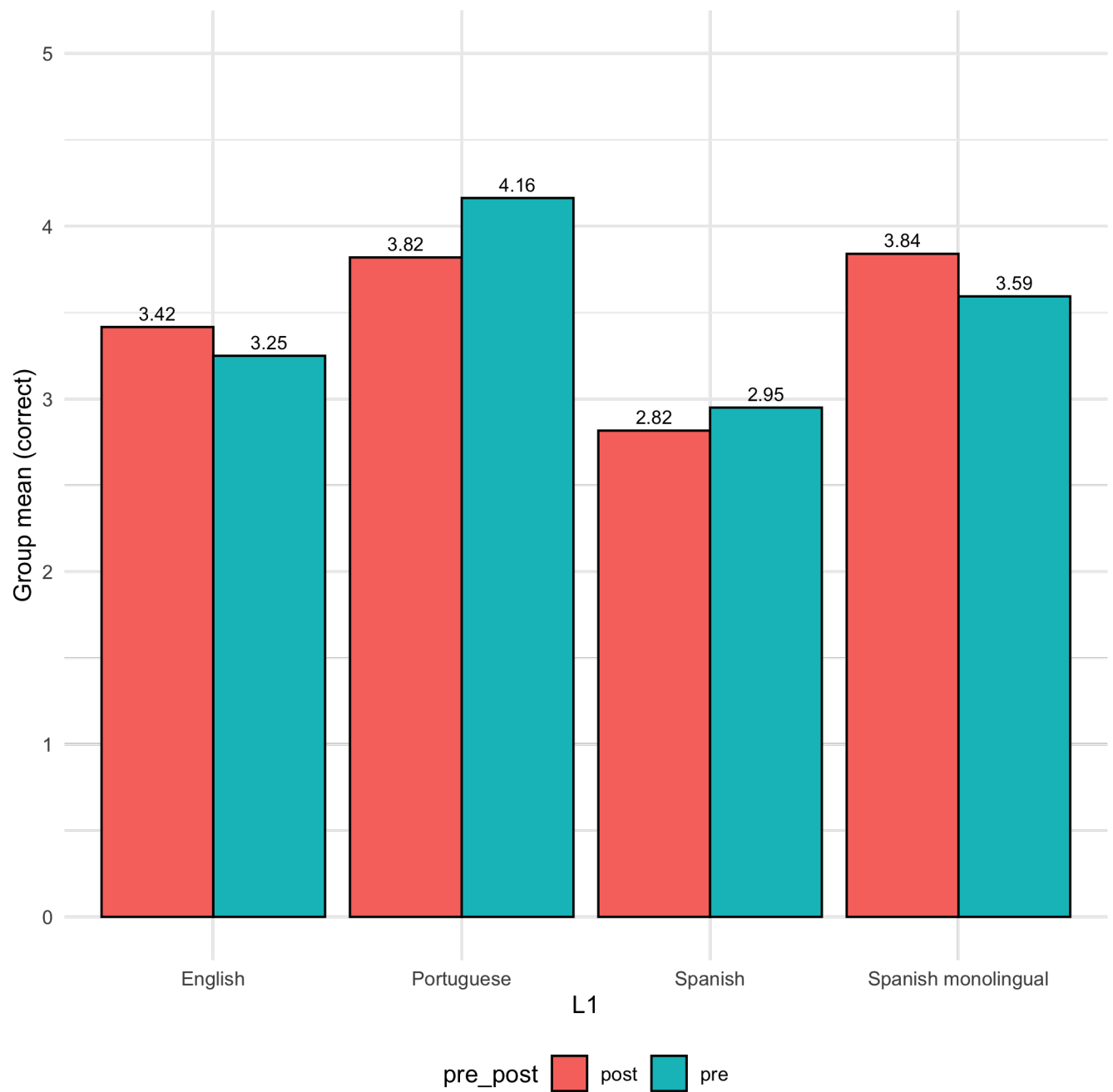


Figure 2. Semantic Interpretation Task

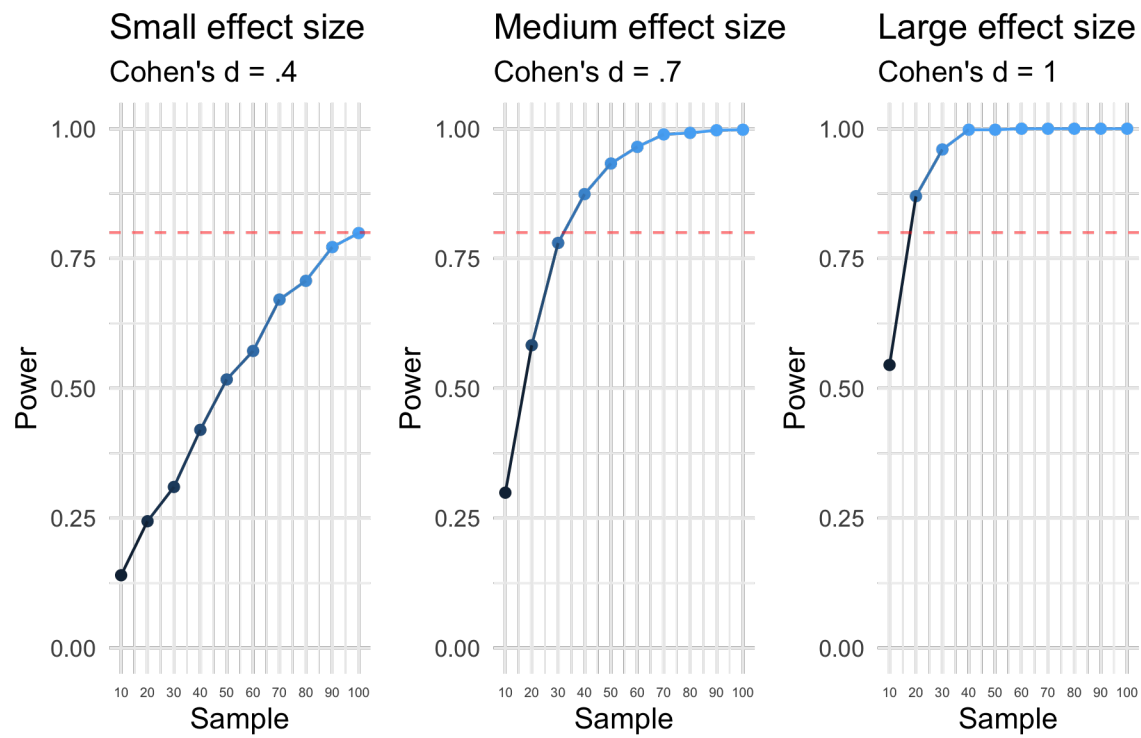


Figure 3. Power Analyses for small, medium and large effect sizes

Results of 2000 simulated t-tests

Replication of Brysbaert (2020)

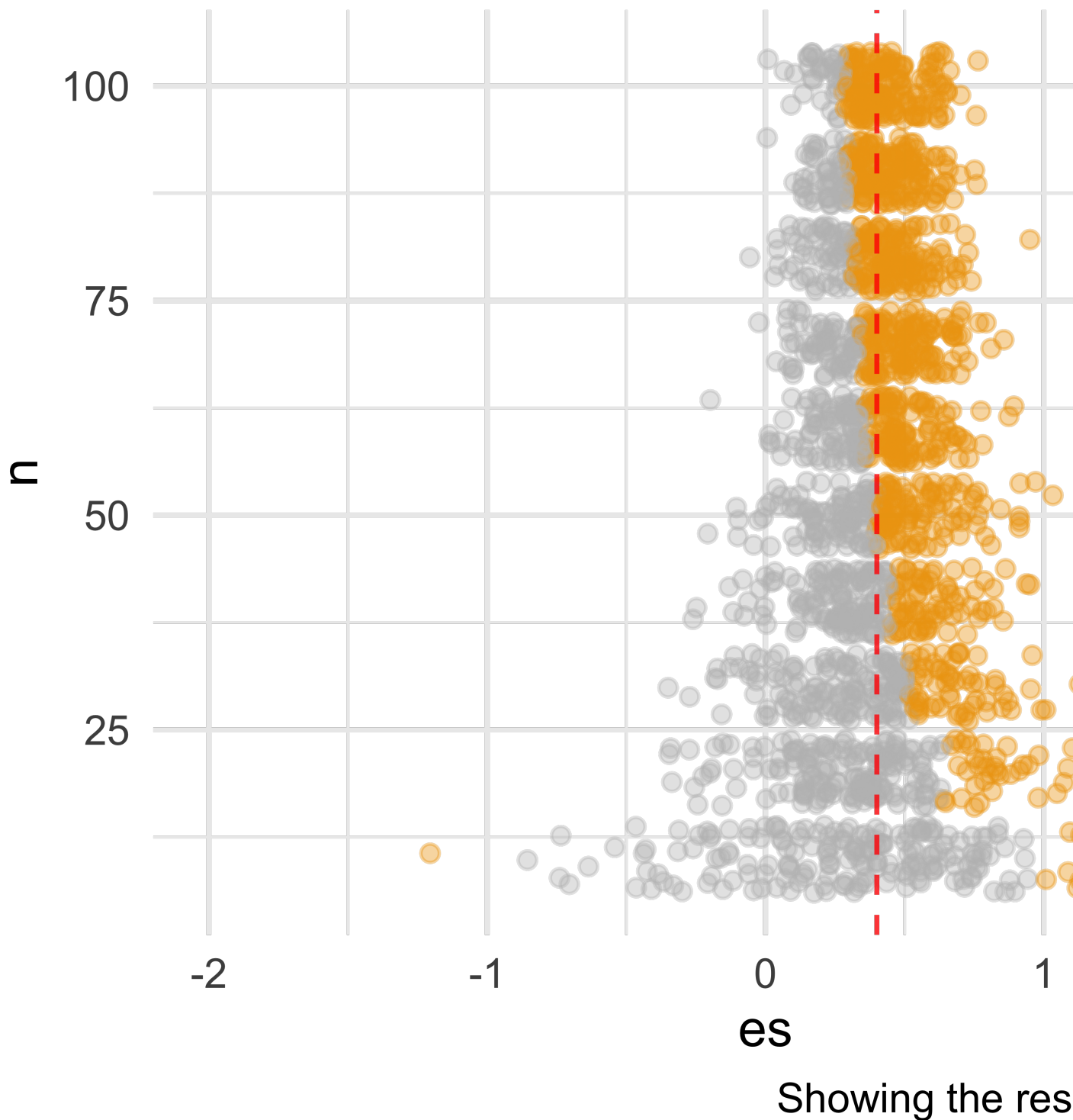


Figure 4. Results from 2000 simulated t-tests