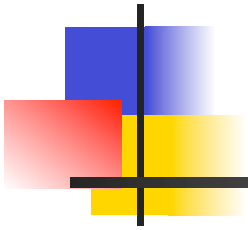


Manuscript Cycles: Coping with rejection and equivocation



Especially equivocation



Trying to see feedback as favorable

- *Your paper received enthusiastic reviews from 3 experts in the field and I am happy to accept it as is.*
- Faint praise: More typical
 - “Enclosed are the reviews of your paper, which were **somewhat favorable**, but also contained **important criticism**. At least one of the reviewers has asked to see a revised version of your paper another time for re-review. **It is possible that this new review will generate additional comments and criticisms that you will also need to address before the paper can be given a final recommendation.** It is usually the case that a reviewer recommends 'revise and resubmit' because she or he sees enough of value and interest in the paper that eventually may render it acceptable for publication. At the same time, he or she sees enough serious problems with the paper that the final recommendation is uncertain.”



Coping with negative feedback

- The natural 3-phase student model

1. Rejection and Anxiety
2. Rejection and Despair
3. Rejection and Suppression

2. Rejection
and Anger



A learned model for success

- RPSP: 4-phase Adaptive Model

1. Rejection and Anxiety
2. Reformulation: Rejection-->Advice
3. Problem Solving
 1. **Careful reading: identify what must be fixed; ignore everything else**
 2. **Recruit co-authors (especially faculty) to help**
 3. Decide what can be fixed and then fix it
 4. Return to letter and reviews and respond to all points
 1. Explain changes in letter
 2. Use letter to explain things not changed
4. Perseverance: Complete, respond, and repeat as necessary



Adapting to different kinds of feedback

1. OK experiments, but basic premise is challenged. Argue or go somewhere else.
2. Clear rejection. Go somewhere else.
3. Equivocation with many, many points to fix
4. Still not good enough feedback



1. The review that challenges your basic premise

(Reviewer 1) The studies may have produced some interesting results for some questions about artificial object – artificial word association (**although this is not so clear, as I explain below**). ...However, it is quite overreaching to use these results to argue in favor of a specific instantiation of perceptual symbol theory. The logic of the studies is roughly that the authors want to know the nature of the perceptual representations that experiential theories claim are used in language experience. Are [they] specific to experienced objects or ...abstracted ...to more schematic representations? Their experiments, they say, support the first alternative. Thus, language comprehension relies on object-specific perceptual representations. **At a broad theory level, this logic is like assuming that the moon is made of cheese and then asking whether this cheese is green or blue. An experiment on the color of cheese may tell you that cheese is blue rather than green, but that does not tell you that the moon is made of cheese.**

Editor's decision: I have now received comments from three reviewers on your submission. Reviewer 2 has identified himself in his review. **As you will see, there is considerable variation in the enthusiasm the reviewers display towards this manuscript, with recommendations ranging from outright reject (Reviewer 1) to publishable after revisions (Reviewers 2 and 3). On this occasion, my opinion is closest to that of Reviewer 1, and I have therefore decided to reject this manuscript.** An important reason for this decision has to do with the theoretical basis of this paper: you basically assume an embodied view, by which language comprehension involves the activation of perceptual information about referents, and that a priming effect from newly learned objects to the pseudowords they have become associated to in learning, confirms this embodied view. With that out of the way, you then wish to investigate the nature of these perceptual representations.

2. The Rejection from Mount Olympus

“...Upon consideration of the comments from the enclosed reviews and their recommended action, **we are unfortunately unable to accept your manuscript for publication in the *Journal of Supreme Excellence***. Due to space limitations, our decision to publish manuscripts is based on several criteria, which **includes** both the quality of the work, and the level of importance of the work. Unfortunately, although both reviewers were generally positive about your work, neither reviewer rated it high enough for consideration for publication.

Please understand that competition for space in *JSE* has become increasingly intense as we receive an ever-growing number of submissions. As a result, *JSE* cannot accommodate all submissions that may ultimately be publishable. Thus, these difficult editorial decisions are becoming increasingly frequent. We hope the enclosed remarks are helpful and will guide you in your future efforts on this work. **Our best wishes as you submit your manuscript elsewhere and thank you for submitting your work to JSE”**





3. The avalanche review with equivocation

Reviewer #1:

In this report ERPs are used to examine "... the hypothesis that the comprehension difficulties of less-skilled comprehenders stem from a semantic processing difficulty." (and not from visual decoding problems). The definition of more or less skilled is operational, based on a median split of the thirty participants studied. ERP recordings were taken from the participants in three tasks: semantic priming task with words (comparing associative, categorically related, and unrelated words), semantic priming with pictures (related versus unrelated), and a rhyming task (rhyming vs nonrhyming words). The P2 and N400 regions of the ERP were sensitive in the semantic-word task (and not in the two other tasks) to differences between skilled and less skilled comprehenders. In particular, the authors emphasize the similarity of the rhyming effect in the two groups. They thus conclude that "these findings provide the first neurophysiological evidence of semantic processing differences between adult skilled and less skilled comprehenders".

While there is much to recommend this study, I also have some concerns.

On the one hand, the authors took great pains to get various behavioral measures in order to better characterize their participants. **On the other hand**, we don't know pre or post grouping by skill, the gender makeup of the groups, handedness, family history of left handedness; I presume that this information is available just not reported. Less likely to be available, but I thought I would ask as I think they would be useful: Was there any span measure of assessment of working memory? Was only one subtest of the Nelson-Denny given and what is the range of possible scores on the test given? Given how much factors are probably uncontrolled and unknown, more information should be given about those that are so we can be more certain that the observed pattern of differences does in fact reflect what the authors claim.



Reviewer 1 cont.

One methodological concern that I have which impact the data has to do with the fact that stimulus duration on the target items (if I understand correctly) was determined by each participant's response - i.e., the stimulus was on until a response was given. This means that it differed across conditions and participants. This makes the ERPs elicited difficult to interpret. The off potential elicited by the end of the stimulus presentation thus could have differed in a systematic fashion across conditions and across groups. We don't know. This information is crucial. I need to be convinced that variance in this is not systematic in a way that could influence the presence or absence of the ERP effects reported.

I must admit I don't understand the point of recording so many channels when there are no good theoretical reasons for so doing, and especially when there is a tendency to then limit analyses to an even smaller subset than those who record a smaller number initially. For example, it is stated that "bad channels were removed from the recordings and replaced with spherical spline interpolation from the remaining channels". Not that I agree with the selection of channels recorded, but even so, were any of the channels on which the analyses were conducted interpolations of this sort in any subject? If so, how many, and which group were they in?

The Method (for selecting electrodes?? Or whatever analyses was performed) from the referenced dissertation needs to be explained in greater detail and justified as it has not gone through peer review process, and it is unreasonable to expect readers to search it down.



Reviewer 1 cont.

Smaller details:

When the authors state "N400 and is typically associated with semantic incongruity", it depends on what is meant by incongruity. Much effort has gone into showing that the N400 is in fact not a sign of incongruity, so maybe the authors would consider using a different term such as unrelated or less related

I have no idea what this sentence means "Therefore, at least in some tasks the N400 seems to be sensitive to differences that are not conscious expectancy violations".

5 An exception was made for a few subjects who had slightly less than 15 trials (11-14) in one or more conditions but whose overall data was very clean. Trial counts of 15 are too small!

When discussing the results of others it might be a good idea to mention which use average reference and which not as this too influences comparisons across data sets.

No interactions with skill on any tasks Σ just main effects Σ in RT

Lower skill slower

I don't always buy into this but it is often argued that early semantic effects such as on the P2 are due to the onset of an N400. It's hard to evaluate such claims when totally different channels are used for analyzing different components. The traditional approach is to do analyses with full head coverage, find interactions with electrode and then do the appropriate follow-up comparisons.



Reviewer 1 continued

It is difficult to see the relevant P200 relatedness effect in Figure 2; the authors might consider putting a box around it, or having a one channel insert blow-up, or changing the scale and doing something to point it out.

Pg. 18 -- ..what does broader Σ refer to specifically; what is broad?

Likewise, on the N400, I would prefer to see more electrodes analyzed, then an anterior-posterior interaction, and so on. I worry that some of the pairwise comparisons are significant because the sites were selected and limited. On pg. 19, there are interpretations of what the interactions (e.g., relatedness by comprehension by hemisphere) without the follow-up analyses required.

To the eye, the pattern of effects really are a matter of degree more than qualitative differences across the two groups; would we have obtained the same results if the median split were at a slightly different median? Is there a cutoff, or would this be seen for any relative split?

Were the subjects debriefed as to how they feel about reading, or how much reading they do, or how they felt about the experiment? Is it possible that the group difference is related to their differential (e.g., emotional) responses to and expectations for their performance in the tasks?

Pg. 25 Although authors end of a note of caution in the paragraph given that they draw such a strong conclusion from a null effect - nothing with pictures so much be limited to verbal domain - I would like to see the caution incorporated with the conclusion - and the entire discussion shortened considerably.



Reviewer 2

Reviewer #2: [The authors] submitted an original study, that examines an interesting question: whether the comprehension problems of adults may be caused by a semantic processing difficulty, suggesting a general link between comprehension skill and semantic skill. To do so they compared 15 skilled comprehenders (SC) and 15 less-skilled comprehenders (LSC) in an EEG study of semantic (compare two words or 2 pictures) and phonological (compare homophones) tasks.

I think the study deserves publication in B&L, **however I have some doubts regarding the findings that I think should be addressed first by the authors before publication.**

Main difficulty:

There is an un-explained difference between the **behavioural** results and the EEG patterns. From the RT results, it is clear that SC were faster than LSC, however the pattern of behaviour was similar - semantic priming was evident only for the associative, not the categorical pairs. In addition, both groups show phonological priming, however again the SC were faster than the LSC.

So we have here 2 groups of subjects, both showing (in RT) similar associative semantic priming, phonological priming and lack of categorical semantic priming. One group is always faster than the other group - so the authors should convince us it is not simply general processing differences, due, for example, to poorer working memory among the LSC.



Reviewer 2 continued

To allow expectancy effects in the EEG study, the task relies on working memory. However, there is evidence that poor comprehenders have difficulties in holding information in working memory (themselves mentioned it on p. 4, Oakhill et al, 2003); So perhaps the poorer performance of LSC reflect working memory functions. **One way to check it would be to repeat the same tasks but when the two stimuli (to be compared whether mean the same or sound the same) are presented together, so the memory component is taken away.** The prediction based on the current study is that poor comprehenders should be poorer on the semantic task when the two items are presented simultaneously and not as a prime and target. They should not differ on the phonological task (no need to do the EEG again, it is only an example of a behavioural task that might explain the RTs differences that were not reflected in the EEG).

The EEG results are interesting however as long as they do not fit the RTs results, there might be alternative accounts. Not only the lack of interaction in the RTs, but also note that the SC have unique EEG patterns in response to categorical priming, which does not reflect in performance at all. This also needs to be discussed. In summary, the authors should elaborate on the relationship between the RTs results (which show only general slower RT among the LSC compared to SC but similar priming effects) and the EEG patterns.

I suggest to do F2 analysis (at least for the behavioural results) to make sure the patterns are not artefact of the specific items used in this study



Reviewer 2 continued

Minor comments

I assume (by the DF) that the analysis was a mixed analysis, with repeated measures for relatedness and subjects group as the between subject variable (not only repeated).

I couldn't figure out what is shown in Table 1 and 2 (they have identical title and headlines) - I assume one is semantic word and one is semantic picture task, but couldn't get it from the text. They do not show the same picture, and in fact there is quite a lot of semantic activity in the less-skilled comprehenders (LSC) which does not seem to differ from the skilled readers. For example, the difference between Associative-categorical in LH is 0.88 for the SC, and significant, while it is 0.78 for the LSC but not significant; for the middle, categorical-unrelated activity is significant 1.92 for the SC, but a non-significant 1.89 for the LSC. Of course we should trust only significant differences, yet it seems that effect size is rather similar among the 2 groups (which actually fits the RTs data).

It would be useful to publish the stimuli, particularly the phonological pairs as it is a new set of stimuli that is not based on previous work.



4. The still-not-good enough review

The authors have made a number of changes in response to reviews of the original submission, including the report an additional experiment, some additional statistics (reported and not reported) and more text. Overall, the revision has strengthened interpretation of the conclusions the authors advance—that children not only show orthographic consistency effects in speech recognition, these processes reflect sublexical processes.

Although the interpretation seems more secure with the additional work, the paper now strikes me as too long and discursive in proportion to its empirical contribution, which is considerable, and its theoretical contribution, which, although interesting, can be only modest given the experiments and their results, which cannot rule out other interpretations convincingly.



The response to the editor

- A mixed message of
 - respect for the reviewers (and editor)
 - apologia
 - explanations of changes
 - rebuttal of points not leading to change
 - confidence that the paper now makes a clear contribution
- All revealing significant effort on your part.

“This revision is fully responsive to your comments and those of the reviewers. In what follows, I describe what we have done to change the paper in response to the criticisms. And, in a few cases, I explain why I think no change was the better course.”

“We appreciate your comments as well as those of the reviewers ... We regret not being as responsive as we should have and we apologize for text problems”

Next slide

Examples are excerpts from an 8 page letter with final revision



Rebuttal of points leading to no change

- “We do not agree that we need to test the model with lexical decisions. Other models of naming seldom do this, and when they do they must resort to non-obvious assumptions about what the input for a decision is. Simulation models are specific in their scope, for better or worse, and ours is a model of primed naming that has implications for character identification in general. We recognize that bringing out these implications when only naming has been used is subject to disagreement. But it is a value of modeling to discuss its broader implications, and await further model development to be corrective if necessary.”



It is possible to argue (but get your advisor to do it for you)

“In that letter, I acknowledged the value of the criticisms of two of the reviewers and objected strongly to a third reviewer, who had some axe to grind...

"There is a fundamental flaw in this claim, and for one of the reviewers (C) an indulgent exercise of using his review to report data of his own. Reviewer C demonstrates merely that you can't find something if you don't look for it. He used the SOAs we reported and failed to find facilitation at the shortest SOA, 43 ms, in contrast to our result. (He ignored the fact that he found a small difference in the right direction.) Our model predicts that there will be a brief period of facilitation based on shared graphic form followed by a period of inhibition as the character reaches its identification threshold. In our published experiment this shift occurred between 43 and 57 ms. In some other lab situation, depending on exact luminance contrast and other display parameters and on individual subject differences, this may vary. If someone were really interested in replicating our result, they would have tried other SOAs... My inference is that Reviewer C has some personal goal here rather than replication. The tone of his review is consistent with this inference. "



Perseverance

- “I think it is in fine shape and I am pleased to accept it for publication in the *JECP*”--