Peer review correspondence

Ms Title: Proximity and Expectancy Modulate Response Vigor After Reward Omission

Author names: Zhang Chen, Christina B. Reimer, Frederick Verbruggen

Submitted: June 22, 2020

Editor decision: Revise & Resubmit

August 4, 2020

Dear Dr. Chen,

I have now received all reviews of your manuscript, “Proximity and Expectancy Modulate Response Vigor After Reward Omission” from qualified researchers. I also independently read the manuscript before consulting these reviews. I agree that your manuscript has important strengths and also that there are some issues that need to be addressed. I therefore encourage you to submit a revised version for further consideration at Collabra: Psychology.

The reviewers did an outstanding job in their reviews and I would like to publicly thank them for devoting their time in carefully commenting on the manuscript. The Reviewers’ comments are stated clearly below so I do not need to iterate them here. As you will see, Reviewer 1 has a series of thought-provoking, theoretical comments. Reviewer 2 comments (mostly) on the used methodology as well as makes many good points for increasing the clarity of the manuscript (e.g., see comment for Figure 1).

From my independent reading, I found that your work is strong, with each experiment building nicely on the findings of the previous one. I want to also compliment you for following open science practices, including the sharing of data and analyses scripts. However, I have some comments as well that I list below:

1. Please comment on the accuracy of measuring RTs online (also raised by Reviewer 1).
2. When you refer to Bayes factors (BFs) on p. 15, it is important to note that Bayes factors quantify the relative evidence of the data coming from the two hypotheses (here H1 and H0). I feel that your text where it is mentioned “to quantify evidence for both the alternative and the null hypothesis” as well as later where you mention that [we used] "BF > 3 to show support for the alternative hypothesis, and BF < 1/3 to show support for the null hypothesis” maybe gives the impression that the BFs provide evidence of the data coming from either the H1 or the H0.
3. Also, for the Bayesian t-tests, I would encourage you to run a sensitivity analysis – it is rather easy to do that in JASP – and monitor whether the direction of the Bayes factors change depending on the value of the scale factor for the Cauchy distribution.
4. It should be noted that the cut-off categories of the BFs that are mentioned in the manuscript are accepted from only part of the literature; BFs can just be used as a continous measure of evidence.

I ask you that in your resubmission, you include a document with a point-by-point response to both the points I list here and the reviewers’ comments, outlining each change made in your manuscript or providing a suitable rebuttal.

In summary, I think this is a promising manuscript and, I hope you will revise it for further consideration at Collabra: Psychology. I look forward to receiving your revision.

Please ensure that your revised files adhere to our author guidelines, and that the files are fully copyedited/proofed prior to upload. Please also ensure that all copyright permissions have been obtained. This is the last opportunity for major editing, therefore please fully check your file prior to re-submission.

If you have any questions or difficulties during this process, please contact the editorial office at [editorialoffice@collabra.org](mailto:editorialoffice@collabra.org).

We hope you can submit your revision within the next six weeks. If you cannot make this deadline, please let us know as early as possible.

Sincerely,

Angelos Krypotos

**Reviewer 1**

**Open response questions**

Please write your review here. The author(s) will see this review. Your identity will not be revealed to the authors unless you also include your name (i.e., sign your review) in this box. It is up to you whether to reveal your identity or not, either is fine.

In this manuscript titled “Proximity and Expectancy Modulate Response Vigor After Reward Omission”, the authors investigate the effect of proximity and expectancy on response vigor (i.e., operationalized as reaction times moving through the trials in a gambling task). In three online studies, they find that participants reacted (confirmed) more quickly after a loss than after a win. Furthermore, they found that proximity and expectancy had different influences on vigor. My overall impression was that this is solid systematic work that exemplarily followed open-sciences practices. Hence, in this sense, I think the findings reported in this manuscript are robust and reliable. However, I have to say that, though I have affinity with research on learning tasks and expectations, the topic of this paper is somewhat far removed from my own research and therefore I found it quite difficult to completely follow the paper.

One thing that struck me though is that the authors consistently differentiate between expectancy and proximity. That I found somewhat puzzling. That is, expectancy is a cognitive event regarding the anticipation of a reward/punishment. Proximity, on the other hand, seems to refer the extent to which the observed chance events (e.g., AAB, ABA, ABB) were close to obtaining a win (e.g., AAA). I found this distinction somewhat strange, because these two constructs are entirely differently operationalized: one is a cognitive event and the other is a property of the procedure. I wonder whether proximity could also not be redescribed as a type of expectancy. For instance, on Page 7, the events ‘ABC’ and ‘ABA’/’ABB’ are described as identical in expectancy, but different in proximity. However, I think these two events differ a lot in terms of expectancies for the *next trial*. ‘ABA’/’ABB’ will likely be interpreted as ‘being close to winning’, while ‘ABC’ is probably not. So I think that (the effects of) proximity could probably be rediscribed in terms of expectancies. As such, I found it somewhat odd to write on Page 45 that “We propose using the separate dimensions of proximity and expectancy to define and operationalize different types of near miss”. Proximity and expectancy are not two entirely different entities in my view. Expectancy is a cognitive event that can be a function of regularities in the environment/the procedure (such as the proximity to obtaining a win).

It is also somewhat strange I find to consider expectancy as rational/following probability theory. ABC and ABA likely do not elicit comparably levels of expectancy (in contrast to what is written on Page 7) because people are not really rational and have a poor grasp of chance events (which we know from the work of, among others, Kahneman and Tversky and phenomena such as the Gambler’s fallacy). I think it would be good if the authors could give some more consideration in the manuscript regarding the distinction between expectancy and proximity. To me at least, it doesn’t entirely make sense as they are situated on two different levels of analysis (i.e., one on a cognitive and the other on a procedural level; see De Houwer, 2011).

Furthermore, I think that part of the struggle that I had with the manuscript is because I found it difficult to follow the theory. The authors mention appraisal theories and I agree that appraisal theories are somewhat relevant for the topic at hand to some extent (i.e., because appraisal theories assume that actions and emotions serve to reach some goals; which are aspects that are also present in the task that the authors use). However, appraisal theories are typically applied to emotion research and this is not really something that the authors consider. Perhaps an alternative/better approach to describe the predictions in terms of utility theory (i.e., that participants’ responses/effort would be a function of their expectations of a reward and the size thereof; <https://en.wikipedia.org/wiki/Utility>) and possibly contrast it with alternatives (e.g., prospect theory by Kahneman and Tversky; <https://en.wikipedia.org/wiki/Prospect_theory>). The prediction could then be possibly then rephrased as whether participants’ follow a utility function, which would be completely rational (e.g., identical responses to ABA and ABC trials; this is what the authors called expectancy I think). Or whether participants’ responses is influenced by irrelevant features from a probability theory perspective such as proximity (and which likely influences participants’ prospects/expectations of being rewarded and hence their vigor on the next trial).

Taken together, I would suggest that the authors try to formulate their predictions in terms of expectations, rather than mixing procedural and cognitive explanations. Furthermore, I think that perhaps economic theories could help with formulating clearer and more straightforward predictions so that the manuscript may be easier to follow. Finally, I think that one nice extension of this task would be to actually *measure* expectations during the task, for example by asking “to what extent did you expect to win a reward” in every 1 out of 5 trials (to not break the flow of the experiment). I think this would likely show that trials that are expected to be equal in expectations (e.g., ABC and ABA) are actually not. Concurrently measuring expectations works quite well within conditioning set-ups (e.g., Boddez et al., 2013).

References Boddez, Y., Baeyens, F., Luyten, L., Vansteenwegen, D., Hermans, D., & Beckers, T. (2013). Rating data are underrated: Validity of US expectancy in human fear conditioning. Journal of behavior therapy and experimental psychiatry, 44(2), 201-206. Houwer, J. D. (2011). Why the cognitive approach in psychology would profit from a functional approach and vice versa. Perspectives on Psychological Science, 6(2), 202-209.

**Rating scale questions**

|  | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| --- | --- | --- | --- | --- | --- |
| The study/studies in this manuscript have strong construct validity (good measures and/or manipulations of the constructs the authors wish to study). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  |  | ✔ |
| The study/studies in this manuscript have strong statistical validity (appropriate statistical tests, assumptions are clear and reasonable, no statistical errors, appropriate statistical inferences, etc.). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  |  | ✔ |
| The study/studies in this manuscript have strong internal validity (any causal claims or implications are well-justified, alternative explanations are thoroughly considered, etc.). (Choose “Neutral” if this is not an empirical manuscript, or no causal claims are made or even vaguely implied.) |  |  |  | ✔ |  |
| The study/studies in this manuscript have strong external validity (authors appropriately constrain their conclusions based on the limits of the generalizability of their findings to other contexts (including from lab to real world), other populations, other stimuli or measures, etc.) |  |  |  | ✔ |  |

**Reviewer 2**

**Open response questions**

Please write your review here. The author(s) will see this review. Your identity will not be revealed to the authors unless you also include your name (i.e., sign your review) in this box. It is up to you whether to reveal your identity or not, either is fine.

This is an interesting manuscript reporting 3 experiments conducted via Prolific, looking at response vigor in the context of reward omission. It builds nicely on some prior work by Verbruggen and colleagues looking at sequential reaction times following gambled wins and losses. By moving to a ‘scratch card’ task in which each trial comprises 3 successive responses (turning over cards), this manuscript is able to distinguish and characterize different forms of reward omission; effectively, different forms of ‘near-misses’. The approach extends a prior study by Bossuyt et al 2014 which distinguished proximity and expectancy mechanisms; this manuscript measures RT/vigor rather than self-report. The paper is well written, the experiments are carefully designed and analysed, and there is strong attention to open science and supplementing frequentist stats with a Bayesian interpretation.

From an experimental perspective, I feel these studies are somewhat unusual (innovative) in inserting the ‘Confirm RT’ response. From my understanding, this is an incidental response that would not have any analogue in real gambling, nor would it be necessary within many cognitive psychology tasks. What it allows is for the authors to ‘bracket’ the processing of the feedback on trial n, from the voluntary initiation of the next trial, in order to extract vigor. As I was reading the paper, I was initially most interested in their Start RT variable, but it becomes clear over the course of the manuscript that the Confirm RT variable best captures the difference between proximity and expectancy. In turn, this highlights the experimental value in taking this extra response in future work.

The introduction does a good job of framing ‘reward omission’ as a neglected topic, and integrating prior work on gambling, behaviourist accounts, and more social psychology accounts via the lens of emotional appraisal. On pg 3, I would reword the assertion that near misses ‘can even lead to pathological gambling’ – pathological gamblers may have a heightened response to near-misses (and probably Sescousse et al 2016 Neuropsychopharmacology is the best evidence for that statement) but this does not mean that near-misses create the disorder.

In using Prolific and jsPsych for online task deployment, can the authors comment on the accuracy of RT measurement, e.g. across different computers / platforms / background apps?

The authors use the ‘AAA’ (AAB, ABC etc) terminology throughout the paper for denoting the configurations. When first describing the Bossuyt et al study on pg 6, it might help the reader to point out that participants win for both AAA and BBB (in some other papers, I believe only AAA denotes a jackpot). Then, when first introducing their own task on pg 9, it would again be useful to say that the letters denote the configurations rather than unique symbols. (This is to say, AAA can be 3 grapes, 3 oranges or 3 strawberries, which all win an equivalent payout).

I raise the last point because in considering the card3 data (and perhaps to some extent, the ConfirmRT data) I found myself wondering to what extent there is a pure cognitive load from processing symbols as a match? (conversely, can non-match combos be immediately ‘disregarded’, at a relatively ‘perceptual’ level? In a couple of places, the authors propose a more top-down account of this effect, of strategic slowing after the AA\*, to choose the final card. Ultimately, because the card3 data are sensitive to win magnitude and win probability (Expt 3), I think my suggestion is unlikely, but it may warrant some consideration. (The ABA condition in Expt 2 may also be relevant to this issue)

Pg 9: in choosing a scratch card over the more standard slot machine design, the authors say that programming spinning reels is difficult. That is true; but it is their choice to use the scratchcard task that allows the 3 successive card RTs to be recorded. (On a slot machine task, this would lack ecological validity). As such, this paper adds to the recent work by Madison Stange on scratchcard near-misses, and shows the value of studying these effects outside of the narrow context of slot machine games.

Fig 1. The text on pg 12 implied to me the outcome screen read the gross gain (60) but the Figure displays the net gain (50)?

In each expt (e.g. pg 13) the authors comment on the balance of wins and losses, and frequency of wins, but they do not formally comment on the overall expectancy of the task. Gambling scenarios classically have a house edge (negative expectancy). Expt 1 has a positive expectancy, whereas Expt 2 (if I understand correctly) has a neutral expectancy? This is then formally manipulated in Expt 3.

P14: “Note that after the second response has been executed on trial n+1, the outcome is partially revealed” – this also applies to trial n, does it not?

Fig 2, 3 and 4. The card3 analysis distinguishes 2 levels of AA\* and AB\*. This seems appropriate as the final card is unresolved; therefore I was not sure about distinguishing all 4 levels in the lefthand panel of the figures? (6 levels in Fig 3).

Pg 21 “rather than disengaging”. I am not sure that the authors can infer in this task which RT direction would constitute ‘disengaging’. In ‘writing off’ a loss trial, the participant may rapidly complete all their remaining responses. Alternatively, unmotivated behaviour or mind wandering could express as the more conventional slowing of RT.

I hope these comments are useful to the authors in preparing this manuscript for publication. Signed, Luke Clark

**Rating scale questions**

|  | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| --- | --- | --- | --- | --- | --- |
| The study/studies in this manuscript have strong construct validity (good measures and/or manipulations of the constructs the authors wish to study). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  |  | ✔ |
| The study/studies in this manuscript have strong statistical validity (appropriate statistical tests, assumptions are clear and reasonable, no statistical errors, appropriate statistical inferences, etc.). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  |  | ✔ |
| The study/studies in this manuscript have strong internal validity (any causal claims or implications are well-justified, alternative explanations are thoroughly considered, etc.). (Choose “Neutral” if this is not an empirical manuscript, or no causal claims are made or even vaguely implied.) |  |  |  |  | ✔ |
| The study/studies in this manuscript have strong external validity (authors appropriately constrain their conclusions based on the limits of the generalizability of their findings to other contexts (including from lab to real world), other populations, other stimuli or measures, etc.) |  |  |  | ✔ |  |

Author response letter:

September 4, 2020

 Dear Dr. Krypotos,

First of all, we would like to thank you and the two reviewers for taking your

time to read our manuscript carefully and providing very thoughtful comments.

Below we list the comments by you and the two reviewers (in italics), and specify

how we have addressed these issues in the revision. In addition to the changes

listed below, we also made some minor textual changes to hopefully make the

manuscript clearer. We think addressing these comments has further improved

the manuscript, and hope that you will be satisfied with the revision. We look

forward to your response to our revised manuscript.

Comments by the editor

1. Please comment on the accuracy of measuring RTs online (also raised by

Reviewer 1).

 We now provide more information on the accuracy of measuring RTs online

with jsPsych.

"Although the response times registered by jsPsych in Chrome and Firefox

have a lag between 23 and 54 milliseconds, the variability across trials is

relatively small (the inter-trial standard deviation of response times caused by

browser/operating system configurations varies between 3.23 and 8.37 milliseconds,

Bridges et al., 2020), and is comparable to other software widely used to

register response times in the laboratory, such as the Psychophysics toolbox (de

Leeuw & Motz, 2016) and E-Prime (Hilbig, 2016). Since we manipulated the

factors of interest within participants, the potential lags in response times (introduced

by different browsers, operating systems, devices and programs running

in the background etc.) should be relatively constant across different conditions

for each participant. The within-subjects comparisons would therefore not be

substantially influenced by the lags in response times." (Page 11-12)

2. When you refer to Bayes factors (BFs) on p. 15, it is important to note

that Bayes factors quantify the relative evidence of the data coming from the

two hypotheses (here H1 and H0). I feel that your text where it is mentioned

“to quantify evidence for both the alternative and the null hypothesis” as well

as later where you mention that [we used] "BF > 3 to show support for the

alternative hypothesis, and BF < 1/3 to show support for the null hypothesis”

maybe gives the impression that the BFs provide evidence of the data coming

from either the H1 or the H0.

 We now provide a formal definition of Bayes factor, to assist the interpretation

of the results.

"...the Bayesian paired-samples t test (to quantify the relative strength of

evidence for two competing hypotheses, Dienes, 2014). A Bayes factor (BF10 )

of B indicates the data are B times more likely under the alternative hypothesis

(the default prior Cauchy’s width = 0.707) than under the null hypothesis."

(Page 18)

1

3. Also, for the Bayesian t-tests, I would encourage you to run a sensitivity

analysis – it is rather easy to do that in JASP – and monitor whether the

direction of the Bayes factors change depending on the value of the scale factor

for the Cauchy distribution.

 We have checked the robustness of the Bayes factors using different priors

in JASP, as you suggested. For the majority of the analyses, the direction of

BF10  did not change (BF10  remained above 1 or below 1). In a few cases where

BF10  did change direction (from below 1 to above 1 when we used increasingly

small prior width), initially inconclusive results remained inconclusive regardless

of which prior width we used (i.e., BF10  remained between 3 and 1

3  ). Using

different priors thus did not change the qualitative conclusions for our results.

We have uploaded the JASP files to the OSF, showing the results of these checks.

In the manuscript, we mention that the BFs are robust against different priors,

and also provide links to the results of robustness checks.

"Furthermore, we conducted robustness checks of BF10  in JASP (0.11.1,

JASP Team, 2019), by using different prior widths. The qualitative conclusions

using 3 and 1

3  as the cut-off values remain the same when using different priors.

The results of these robustness checks can be found at https://osf.io/erhft/,

https://osf.io/a6kep/, and https://osf.io/fxjhn/, for the three experiments respectively."

(Page 18)

4. It should be noted that the cut-off categories of the BFs that are mentioned

in the manuscript are accepted from only part of the literature; BFs can just be

used as a continous measure of evidence.

 We now add such a statement when introducing cut-off values for the BFs.

"While the BF10  is a continuous measure of the relative likelihood of data

under two hypotheses, we adopted the conventional cut-off values of 3 and 1

3  as

substantial evidence for the alternative and the null, respectively, to facilitate

statistical inference (Dienes, 2014)." (Page 18)

Comments by Reviewer 1

1. My overall impression was that this is solid systematic work that exemplarily

followed open-sciences practices. Hence, in this sense, I think the findings reported

in this manuscript are robust and reliable. However, I have to say that,

though I have affinity with research on learning tasks and expectations, the topic

of this paper is somewhat far removed from my own research and therefore I

found it quite difficult to completely follow the paper.

 We thank Reviewer 1 for these positive comments. We have tried to add

more clarifications to our theoretical framework, and hope that our revisions

have made this paper easier to follow for a broader audience.

2

2. One thing that struck me though is that the authors consistently differentiate

between expectancy and proximity. That I found somewhat puzzling. That is, expectancy

is a cognitive event regarding the anticipation of a reward/punishment.

Proximity, on the other hand, seems to refer the extent to which the observed

chance events (e.g., AAB, ABA, ABB) were close to obtaining a win (e.g.,

AAA). I found this distinction somewhat strange, because these two constructs

are entirely differently operationalized: one is a cognitive event and the other

is a property of the procedure. I wonder whether proximity could also not be

redescribed as a type of expectancy. For instance, on Page 7, the events ‘ABC’

and ‘ABA’/’ABB’ are described as identical in expectancy, but different in proximity.

However, I think these two events differ a lot in terms of expectancies

for the next trial. ‘ABA’/’ABB’ will likely be interpreted as ‘being close to winning’,

while ‘ABC’ is probably not. So I think that (the effects of ) proximity

could probably be rediscribed in terms of expectancies. As such, I found it somewhat

odd to write on Page 45 that “We propose using the separate dimensions

of proximity and expectancy to define and operationalize different types of near

miss”. Proximity and expectancy are not two entirely different entities in my

view. Expectancy is a cognitive event that can be a function of regularities in

the environment/the procedure (such as the proximity to obtaining a win).

 We agree that our initial definitions of expectancy and proximity were (too)

brief. In the revision, we now discuss the two separate processes underlying

expectancy and proximity in more details (in the context of slot machine gambling),

and also provide a clarification to what we (and Bossuyt et al.) mean

by expectancy.

"Two unique cognitive processes may be triggered by the unfolding events

in a ’near miss’, which may distinguish it from other types of losses. Take

again a three-reel slot machine as an example. Players may start a game with a

certain expectation of winning. This initial expectation may be influenced by the

outcomes of previous games, their overall optimism, experience with gambling,

etc. After starting a game, their expectation of winning would presumably

increase after seeing the jackpot symbols on the first two reels. When the

third reel eventually stops with no jackpot symbol (creating a ’near miss’),

the loss may be more unexpected compared to a ’full miss’ (where the first

or second reel already shows no jackpot symbol). In addition to an increased

expectation of winning as the game evolves, a ’near miss’ is also subjectively

closer to a win compared to a ’full miss’ (although objectively a ’near miss’ is

as proximal to a win as any other types of losses, M. R. Dixon & Schreiber,

2004), as it contains more jackpot symbols than a ’full miss’. Bossuyt et al.

(2014) distinguished between these two processes, and referred to them as the

appraisals of ’expectancy’ and ’proximity’, respectively. We will adopt this

terminology in the present research for the sake of consistency. Expectancy

and proximity thus describe how the players appraise the unfolding events at

different stages as a game evolves. Expectancy refers to whether the players’

expectations of winning increase from the beginning of a game to the moment

before the outcome is revealed, whereas proximity refers to whether the players

3

subjectively perceive a loss to be proximal to a win, after the eventual outcome

is revealed. Different outcomes may also influence the initial expectation of

winning for the next game, but this is not what we (and Bossuyt et al.) studied

here." (Page 5-6)

Thus, both expectancy and proximity refer to cognitive processes that are

triggered by unfolding events as a game evolves. By expectancy, we (and

Bossuyt et al.) specifically refer to the changes in expectancy of winning from

the beginning of a game till the moment before the outcome is revealed. From

this perspective, the events ‘ABC’ and ‘ABA’/’ABB’ are identical in expectancy

(using our and Bossuyt et al.’s definition), since right before the last card is revealed,

players essentially get the same card sequence in both cases (AB\*),

and their expectation of winning would similarly decrease (as the game unfolds)

in both cases. Note that by expectancy, we are not referring to the

expectation of winning for the next game, which may indeed be influenced by

the outcome of the current game, as the reviewer suggested. To better convey

the idea that by ’expectancy’ we are referring to changes in expectancy of

winning as a game evolves, we now consistently refer to AAB as a condition

with increased  expectancy, and ABA/ABB/ABC as conditions with reduced

 expectancy throughout the whole manuscript.

3. It is also somewhat strange I find to consider expectancy as rational/following

probability theory. ABC and ABA likely do not elicit comparably levels of expectancy

(in contrast to what is written on Page 7) because people are not really

rational and have a poor grasp of chance events (which we know from the work

of, among others, Kahneman and Tversky and phenomena such as the Gambler’s

fallacy). I think it would be good if the authors could give some more

consideration in the manuscript regarding the distinction between expectancy

and proximity. To me at least, it doesn’t entirely make sense as they are situated

on two different levels of analysis (i.e., one on a cognitive and the other on

a procedural level; see De Houwer, 2011).

 We have now provided more clarifications to our theoretical framework, especially

to what we mean by expectancy (see above). ABC and ABA may thus

indeed lead to different expectations of winning for the next game, but that is

not what we (and Bossuyt et al.) mean by expectancy here.

4. Furthermore, I think that part of the struggle that I had with the manuscript

is because I found it difficult to follow the theory. The authors mention appraisal

theories and I agree that appraisal theories are somewhat relevant for the topic

at hand to some extent (i.e., because appraisal theories assume that actions and

emotions serve to reach some goals; which are aspects that are also present in the

task that the authors use). However, appraisal theories are typically applied to

emotion research and this is not really something that the authors consider. Perhaps

an alternative/better approach to describe the predictions in terms of utility

theory (i.e., that participants’ responses/effort would be a function of their expectations

of a reward and the size thereof; https://en.wikipedia.org/wiki/Utility)

 4

and possibly contrast it with alternatives (e.g., prospect theory by Kahneman and

Tversky; https://en.wikipedia.org/ wiki/Prospect\_theory). The prediction could

then be possibly then rephrased as whether participants’ follow a utility function,

which would be completely rational (e.g., identical responses to ABA and

ABC trials; this is what the authors called expectancy I think). Or whether

participants’ responses is influenced by irrelevant features from a probability

theory perspective such as proximity (and which likely influences participants’

prospects/expectations of being rewarded and hence their vigor on the next trial).

 We think appraisal theories provide a useful framework for our current research.

As we mentioned, according to at least some appraisal theories of emotion,

emotions are multicomponential processes that involve changes in various

components triggered by the appraisal of an event. In the current research we

focus on the motivational action tendency component, which is an important

component of emotions (at least according to some leading theories). Frijda’s

framework (which is our main theoretical framework) even argued that motivational

action tendencies form the key to what we call "emotions". Thus,

although our work does not address the subjective feelings component of emotions,

it does focus on an important aspect of emotions. Furthermore, our

experimental procedure was largely inspired by that of Bossuyt and colleagues,

who also adopted appraisal theories as their main theoretical framework. Our

finding that proximity and expectancy have separate influences on response

vigor parallels their finding that proximity and expectancy have separate influences

on subjective feelings. Thus, appraisal theories of emotion provide a

useful framework for understanding both response vigor and negative emotions

after reward omission.

To make the relevance of appraisal theories clearer, we have revised the part

where we introduced the appraisal theories of emotion, to give more emphasis

to motivational action tendencies and Frijda’s idea that such action tendencies

form key to what we call "emotions".

"Appraisal theories of emotion provide a useful framework for understanding

both response invigoration and the associated negative emotions after reward

omission. According to appraisal theories of emotion, emotions are multicomponential

processes that involve changes in various components, such as appraisal

of an event, motivational action tendencies, physiological reactions, expressive

behaviors (facial, vocal and gestural) and subjective feelings (Moors et al., 2013).

Appraisal refers to the process of assessing and evaluating aspects of an event

that are of significance to one’s well-being (Moors et al., 2013). The most central

appraisal is whether an event is promoting or obstructing one’s goal(s).

Failing to obtain rewards can thus be appraised as a goal-incongruent event,

which can lead to certain action tendencies, such as re-engaging and attempting

to obtain rewards again (other action tendencies are also possible, such as

discarding current reward pursuit, but here we will focus on re-engaging after

reward omission). According to Frijda et al. (2014), these action tendencies,

or states of action readiness, form the key to what we call "emotions". Action

tendencies can manifest in actions with a certain strength and urgency (i.e.,

5

response vigor). Actions are considered "impulsive" when they are triggered by

appraisals and not preceded by (much) deliberation. The strength or urgency

of an (impulsive) action depends on the appraised importance of a goal and the

discrepancy between the current state and the desired state (Frijda, 2010). An

action becomes more vigorous when a goal is perceived to be more important,

and/or when the appraised discrepancy between the current and the desired

state becomes larger. Since failing to obtain rewards often entails a larger discrepancy

than obtaining rewards, response thus becomes more vigorous after

reward omission. In addition to changes in response vigor, changes in physiological

responses (e.g., increase in corticosteroid level, Papini & Dudley, 1997;

and potential increase in skin conductance response, Otis & Ley, 1993; but see

Lole et al., 2012) and expressive behaviors (e.g., frowning after losing in gambling;

Wu et al., 2015) can also occur after reward omission. The integrated and

synchronized changes in all components may lead to specific subjective feelings

that can be categorized and verbalized, giving rise to discrete emotions such as

’frustration’, ’disappointment’, or ’regret’ as its end product (Scherer & Moors,

2019)." (Page 3-4)

We thank the reviewer for sharing the thoughts on the utility theory and

Prospect Theory. We agree Prospect Theory could be relevant, in the sense

that the idea that response vigor is influenced by the discrepancy between the

current and the desired state, rather than the current state alone, echos the

ideas of Prospect Theory (Kahneman & Tversky, 1979). In Prospect Theory,

the value of an option is defined by its deviation from a reference point, rather

than its absolute value per se . Initially developed for decision making under

risk, Prospect Theory has been successfully extended to other psychological

processes, such as goal pursuit. In goal pursuit, goals can similarly be seen as

reference points (Heath et al., 1999), and people may adjust their motivation

based on the perceived deviation between the current progress and their goal

(similar to the desired state in Frijda’s framework).

While we can certainly see the relevance of Prospect Theory to the present

work, we choose to not reformulate our theoretical predictions in terms of

Prospect Theory, for several reasons. First of all, our interest in this project

started with the ’near miss’ effect in gambling. As mentioned in the manuscript,

gambling researchers have proposed different theoretical accounts for ’near miss’

effects, emphasizing either its proximity to a win, or the increased expectation

of winning right before the eventual loss (i.e., expectancy). Distinguishing between

proximity and expectancy would therefore be informative for research

on ’near miss’ effects. Second, our procedure was largely inspired by that of

Bossuyt and colleagues. By showing the separate influences of proximity and

expectancy on response vigor, our findings thus complement those by Bossuyt

et al., and show the utility of appraisal theories in understanding the effects of

proximity and expectancy on different components of emotions. Reformulating

the predictions in terms of Prospect Theory would make the connections

between our work and that of Bossuyt et al. less clear. As we understand

it, Prospect Theory was originally proposed to explain decision making under

6

risk, and thus does not make specific predictions on response vigor. Frijda’s

framework on the other hand makes specific predictions on how response vigor

varies as a function of the perceived discrepancy between the current state and

a desired state, and the effects of proximity and expectancy we observed can

also be easily incorporated into this framework. Thus, we think using Frijda’s

theoretical framework (appraisal theories) provides a more accurate account of

how we arrived at our research questions, makes the connection between our

work and previous work clearer, and also provides a more specific explanation

of the results than Prospect Theory does.

5. Taken together, I would suggest that the authors try to formulate their predictions

in terms of expectations, rather than mixing procedural and cognitive explanations.

Furthermore, I think that perhaps economic theories could help with

formulating clearer and more straightforward predictions so that the manuscript

may be easier to follow. Finally, I think that one nice extension of this task

would be to actually measure expectations during the task, for example by asking

“to what extent did you expect to win a reward” in every 1 out of 5 trials (to

not break the flow of the experiment). I think this would likely show that trials

that are expected to be equal in expectations (e.g., ABC and ABA) are actually

not. Concurrently measuring expectations works quite well within conditioning

set-ups (e.g., Boddez et al., 2013).

 We thank the reviewer again for suggesting these interesting ideas. We hope

our clarifications to the theoretical framework (especially to the concept of expectancy)

have made it clearer that by expectancy and proximity, we mean two

cognitive processes that are triggered by the unfolding events as a game evolves.

We agree with the reviewer that if we were to measure players’ expectations of

winning after ABC and ABA for the next game, players would presumably report

a higher expectation of winning after ABA than after ABC. Proximity of

the current outcome may thus indeed increase the expectation of winning for

the next game, which can be why a proximal loss is perceived to be more desirable

than a distal loss (we now also mention this idea in the general discussion).

However, this is not what we mean by expectancy here. ABC and ABA have

the same expectancy (using our and Bossuyt et al.’s definition), in the sense

that in both cases the expectancy of winning similarly decreased as the game

unfolded, before the last card was turned. We are thus focusing on how changes

in expectation as a game evolves (i.e., increased or reduced expectancy) influences

response vigor, and hope that we have successfully clarified this point in

the revision.

Comments by Reviewer 2

1. The introduction does a good job of framing ‘reward omission’ as a neglected

topic, and integrating prior work on gambling, behaviourist accounts, and more

 7

social psychology accounts via the lens of emotional appraisal. On pg 3, I would

reword the assertion that near misses ‘can even lead to pathological gambling’

– pathological gamblers may have a heightened response to near-misses (and

probably Sescousse et al 2016 Neuropsychopharmacology is the best evidence for

that statement) but this does not mean that near-misses create the disorder.

 We have now reworded the statement, and also cited the mentioned paper.

"Furthermore, players with gambling disorder show amplified responses in

the striatum (a core region of the brain reward circuitry) to ’near misses’ than

players without gambling disorder (Chase & Clark, 2010; Habib & Dixon, 2010;

Sescousse et al., 2016). The latter finding suggests ’near misses’ might contribute

to gambling disorder, although the causal relationship between gambling

disorder and reactivity to ’near misses’ remains unclear." (Page 5)

2. In using Prolific and jsPsych for online task deployment, can the authors

comment on the accuracy of RT measurement, e.g. across different computers

/ platforms / background apps?

 Please see our reply to comment 1 by the editor.

3. The authors use the ‘AAA’ (AAB, ABC etc) terminology throughout the

paper for denoting the configurations. When first describing the Bossuyt et al

study on pg 6, it might help the reader to point out that participants win for both

AAA and BBB (in some other papers, I believe only AAA denotes a jackpot).

Then, when first introducing their own task on pg 9, it would again be useful to

say that the letters denote the configurations rather than unique symbols. (This

is to say, AAA can be 3 grapes, 3 oranges or 3 strawberries, which all win an

equivalent payout).

 When first introducing the Bossuyt et al study, we now made it clearer that

participants won points when three reels had the same fruit picture, by giving

examples of win configurations from Bossuyt et al. "(e.g., three lemons, three

prunes or three melons; AAA)" (Page 7). Furthermore, when first introducing

our own task, we added that "Note that we used three cards with different fruits

(oranges, strawberries, and grapes), and each card was randomly assigned the

role of A, B and C on each trial. AAA, AAB, ABB and ABC thus denoted the

overall configurations rather than specific cards (i.e., AAA could be 3 oranges,

3 strawberries, or 3 grapes)." (Page 10)

4. I raise the last point because in considering the card3 data (and perhaps

to some extent, the ConfirmRT data) I found myself wondering to what extent

there is a pure cognitive load from processing symbols as a match? (conversely,

can non-match combos be immediately ‘disregarded’, at a relatively ‘perceptual’

level? In a couple of places, the authors propose a more top-down account of this

effect, of strategic slowing after the AA\*, to choose the final card. Ultimately,

because the card3 data are sensitive to win magnitude and win probability (Expt

 8

3), I think my suggestion is unlikely, but it may warrant some consideration.

(The ABA condition in Expt 2 may also be relevant to this issue)

 When searching the literature on perceptual matching, we found that an

opposite effect was often observed in simple perceptual matching tasks. That

is, participants tend to match pairs of same stimuli more quickly  than different

stimuli, opposite to what the reviewer suggested here. The difference in

card3 RT after AA\* and AB\* therefore cannot by explained simply by perceptual

matching processes. However, we agree that our previous conclusion that

this effect reflects more top-down, strategic thinking may be unwarranted. Importantly,

regardless of the underlying process(es) for this difference in card3

RT between AA\* and AB\*, it shows that as we expected, participants paid

attention to each turned card as the game evolved, which was an important

prerequisite for our manipulation of expectancy. We now emphasize this in the

revised manuscript, and refrain from interpreting this effect of card3 RT solely

as under the strategic, top-down control of participants.

"For card3 RT, only the main effect of outcome was statistically significant

(Table 1). Participants turned the third card more quickly when the first two

cards differed (i.e., AB\*) in comparison to when the first two cards matched

(i.e., AA\*). This effect cannot be easily explained by perceptual processes (e.g.,

participants identified AB\* more quickly than AA\*), as previous work on simple

perceptual-matching tasks indicates that people tend to match identical pairs of

stimuli faster than different pairs (Goulet & Cousineau, 2020; Nickerson, 1967),

opposite to what we have observed here. Instead, this relative speeding up after

AB\* than after AA\* may be explained by (potential) wins and losses. Response

vigor might increase after AB\* (as AB\* indicated a certain loss), consistent with

the ’increased vigor after a loss’ hypothesis. Alternatively, participants might

strategically slow down after AA\* to think about which card to turn next.

These processes are not necessarily mutually exclusive, and may simultaneously

contribute to the difference in card3 RT after AA\* versus AB\*. Even though

our data do not allow us to disentangle these processes, this difference in card3

RT after AA\* versus AB\* shows that participants were paying attention to each

turned card as they went through a game. This is an important prerequisite

for our manipulation of expectancy, in which we assumed that participants

paid attention to each turned card and adjusted their expectation of winning

accordingly as a game evolved." (Page 19-20)

5. Pg 9: in choosing a scratch card over the more standard slot machine design,

the authors say that programming spinning reels is difficult. That is true; but

it is their choice to use the scratchcard task that allows the 3 successive card

RTs to be recorded. (On a slot machine task, this would lack ecological validity).

As such, this paper adds to the recent work by Madison Stange on scratchcard

near-misses, and shows the value of studying these effects outside of the narrow

context of slot machine games.

 Thank you for sharing these thoughts. Indeed, although initially the main

reason for using a ’scratch card’ task was pragmatic, our task allowed us to

9

register the response time of turning each card. We have now revised the text

to include this reason as well:

"We opted to use a ’scratch card’ task rather than a simulated slot machine

for a few reasons. First, pragmatically, a ’scratch card’ task was easier to program

than the spinning reels in a simulated slot machine. Second, by using a

’scratch card’ task, we were able to explore whether the response time of turning

each card in a game was influenced by the cards turned thus far. Although this

is also possible with slot machines (e.g., press a button to start the spin of a reel

one by one), it would be less ecologically valid. Furthermore, instant scratch

cards share many structure characteristics with slot machines (including ’near

misses’), which led some researchers to propose that "scratch cards are essentially

slot machines on paper" (Ariyabuddhiphongs, 2011). While the structural

characteristics of slot machines and their potential influences on gambling have

been extensively studied, our work adds to some recent work that extends this

line of investigation to other forms of gambling, such as scratch card games

(Stange et al., 2017; Stange et al., 2016)." (Page 9)

6. Fig 1. The text on pg 12 implied to me the outcome screen read the gross

gain (60) but the Figure displays the net gain (50)?

 Thank you for spotting this mistake. Participants indeed saw the gross gain

(60 in this case) rather than the net gain. The figure has been corrected.

7. In each expt (e.g. pg 13) the authors comment on the balance of wins and

losses, and frequency of wins, but they do not formally comment on the overall

expectancy of the task. Gambling scenarios classically have a house edge

(negative expectancy). Expt 1 has a positive expectancy, whereas Expt 2 (if I

understand correctly) has a neutral expectancy? This is then formally manipulated

in Expt 3.

 We now comment on the overall expected value of a game in each experiment.

Note that Experiment 2 also has a positive expected value, as we reduced the

wager amount. We now make this point clear.

For Experiment 1:

"The overall expected value of a game was thus positive in Experiment 1,

in that all participants won 720 points in total. This differs from most real-life

gambling scenarios, in which the expected value of gambling is negative so that

players lose money in the long run." (Page 15)

For Experiment 2:

"To make sure participants would win the same amount of points (i.e., 770

points including the initial 50 points), we reduced the wager such that one lowamount

game cost 1 point and one high-amount game cost 5 points (as opposed

to 2 and 10 points in Experiment 1). The amount of points won for each AAA

outcome remained the same (12 and 60 points). The overall expected value of

a game in Experiment 2 was thus again positive (i.e., participants won points

in the long run)." (Page 26)

10

For Experiment 3:

"The expected value of a game in both the high and medium win probability

’casinos’ was positive (participants won 600 and 180 points, respectively), while

the expected value of a game in the low win probability ’casino’ was negative

(participants lost 120 points in total)." (Page 34)

8. P14: “Note that after the second response has been executed on trial n+1,

the outcome is partially revealed” – this also applies to trial n, does it not?

 This indeed applies to every trial. We have now revised this part.

"Note that after the second card has been turned, the outcome of a trial is

partially revealed. We thus did not include the RT of turning the third card on

trial n+1, as this RT may be influenced by both the outcome of trial n and the

partial outcome of trial n+1, and we did not have enough trials to examine this

interaction effect." (Page 16)

9. Fig 2, 3 and 4. The card3 analysis distinguishes 2 levels of AA\* and AB\*.

This seems appropriate as the final card is unresolved; therefore I was not sure

about distinguishing all 4 levels in the lefthand panel of the figures? (6 levels in

Fig 3).

 We distinguished all different levels for card3 RT in the figures, as a sort

of ’sanity’ check. As the third card is unresolved for card3 RTs, card3 RTs

for AAA/AAB/ABC should be similar, while card3 RTs for ABA/ABB/ABC

should also be similar. As can be seen from the figures, this is indeed the case.

By plotting all conditions for card3 RT separately, we think readers may more

easily see how the RTs for each condition developed across the three panels, as

the RTs for each of the outcomes would be based on the same sets of trials.

To make it clearer that for card3 RT, the last card is still unresolved, we

now add a brief explanation to all figure captions.

For Experiments 1 and 3:

"For card3 RT, the third card is unresolved: AAA, AAB = AA\*; ABB, ABC

= AB\*."

For Experiment 2:

"For card3 RT, the third card is unresolved: AAA, AAB, AAC = AA\*;

ABA, ABB, ABC = AB\*."

10. Pg 21 “rather than disengaging”. I am not sure that the authors can infer

in this task which RT direction would constitute ‘disengaging’. In ‘writing off’

a loss trial, the participant may rapidly complete all their remaining responses.

Alternatively, unmotivated behaviour or mind wandering could express as the

more conventional slowing of RT.

 Indeed, we have now revised this statement as follow:

"This effect suggested that participants still paid attention to the last card,

even after they already knew after turning the first two mismatching cards that

they had lost (i.e., AB\*)." (Page 24)

11

**Editor final decision: Accept**

**October 20, 2020**

Dear Zhang Chen,

I have now had a chance to read over your manuscript “Proximity and Expectancy Modulate Response Vigor After Reward Omission”, along with the letter describing the changes you made.

Thank you for your responsiveness to the concerns that the reviewers and I raised. I want to also warmly thank both Reviewers who agreed to review your manuscript one more time. One of the Reviewer suggests that you rewrite parts of the paper so that it can be more accessible to a wider audience. Although I think that such revisions could be helpful, I think that the paper as it stands now presents both the relevant theories and the results in an excellent manner. As such, I am happy to say that your paper is now officially accepted for publication in Collabra: Psychology. Congratulations. I think your work will make an important contribution to the literature and I look forward to seeing it published. I hope your experiences with Collabra: Psychology have been positive and that you will continue to consider it as an outlet for your work.

You do not have to complete any tasks at this point. Our managing editor will contact you in case there are any pre-prodution file related questions. You will have an opportunity to check the page proofs before we publish your article. Thank you again for publishing in Collabra: Psychology.

Sincerely, Angelos Krypotos

**Reviewer 1**

**Open response questions**

Please write your review here. The author(s) will see this review. Your identity will not be revealed to the authors unless you also include your name (i.e., sign your review) in this box. It is up to you whether to reveal your identity or not, either is fine.

I think that the authors gave my comments adequate consideration. That said, I think it would be very useful for our field to move towards more field-general theories and terms, rather than the now more conventional domain-specific approach. That is, like I indicated in my previous, there is much overlap between the focus of this paper and relevant concepts and theories in, for instance, economy and decision making. Making this connection may open up this research to a wider audience. However, I can understand that this goes beyond the aims of this paper, so I am happy to recommend this work as it currently stands.

**Rating scale questions**

|  | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| --- | --- | --- | --- | --- | --- |
| The study/studies in this manuscript have strong construct validity (good measures and/or manipulations of the constructs the authors wish to study). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  | ✔ |  |
| The study/studies in this manuscript have strong statistical validity (appropriate statistical tests, assumptions are clear and reasonable, no statistical errors, appropriate statistical inferences, etc.). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  | ✔ |  |
| The study/studies in this manuscript have strong internal validity (any causal claims or implications are well-justified, alternative explanations are thoroughly considered, etc.). (Choose “Neutral” if this is not an empirical manuscript, or no causal claims are made or even vaguely implied.) |  |  |  | ✔ |  |
| The study/studies in this manuscript have strong external validity (authors appropriately constrain their conclusions based on the limits of the generalizability of their findings to other contexts (including from lab to real world), other populations, other stimuli or measures, etc.) |  | ✔ |  |  |  |

**Reviewer 2**

**Open response questions**

Please write your review here. The author(s) will see this review. Your identity will not be revealed to the authors unless you also include your name (i.e., sign your review) in this box. It is up to you whether to reveal your identity or not, either is fine.

I was pleased to see this resubmission and I enjoyed engaging again with these fascinating data. The authors have engaged thoroughly with the first reviews, and my comments have been satisfactorily addressed. I look forward to seeing this paper in print. Signed, Luke Clark

**Rating scale questions**

|  | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| --- | --- | --- | --- | --- | --- |
| The study/studies in this manuscript have strong construct validity (good measures and/or manipulations of the constructs the authors wish to study). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  |  | ✔ |
| The study/studies in this manuscript have strong statistical validity (appropriate statistical tests, assumptions are clear and reasonable, no statistical errors, appropriate statistical inferences, etc.). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  |  | ✔ |
| The study/studies in this manuscript have strong internal validity (any causal claims or implications are well-justified, alternative explanations are thoroughly considered, etc.). (Choose “Neutral” if this is not an empirical manuscript, or no causal claims are made or even vaguely implied.) |  |  |  |  | ✔ |
| The study/studies in this manuscript have strong external validity (authors appropriately constrain their conclusions based on the limits of the generalizability of their findings to other contexts (including from lab to real world), other populations, other stimuli or measures, etc.) |  |  |  |  | ✔ |