**Peer Review Communication**

**Ms Title: Economic games: An introduction and guide for research**

**Corresponding Author Name: Isabel Thielmann**

**Submitted: Aug 16, 2020**

**Editor first decision—Revise & Resubmit**

Nov 6, 2020

Dear Dr. Thielmann,

I have now received the review of your manuscript, “Economic games: Modeling, measuring, and understanding prosocial behavior” from a qualified researcher. I also independently read the manuscript before consulting these review and again afterwards. I agree that your manuscript has potential to make a good contribution 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.

I would like to thank the authors for their patience. It took me considerable time to find an expert with the time to provide a review for this paper. And I thank the reviewer for their time and effort in sharing their expertise and in providing constructive feedback for the manuscript. The reviewer did an outstanding job in their review.

I will highlight issues I think are particularly salient here. In your resubmission, please 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. Please highlight the changes in the manuscript (either with track changes or by highlighting them in the response letter). My comments follow now:

1. The reviewer notes that the intended audience isn’t clear. Is the paper aimed at theoreticians or naïve experimentalist social psychologists? After having read the manuscript twice, once before sending out for review and once after receiving the review, my impression is that the paper is aimed at the latter (i.e., naïve experimental psychologists). Of course, I may be wrong. However, if this is the case then there’s a lot of assumed knowledge. I recommend that the authors make clear the intended audience and the necessary adjustments to the manuscript (as per the reviewer’s recommendations).
2. When I first read the paper my thought was that the biggest contribution of the paper is best expressed by the following quote from the concluding remarks: “from observing a player’s behavior in a single game, it is rarely possible to conclusively determine a particular underlying process. However, a specific elegance of games is that one can combine different games or game variants to isolate certain processes and social motives […] Overall, combining different games and/or game variants – and comparing corresponding behaviors – will often provide more fine-grained insights into the psychological processes underlying players’ choices and thereby help illuminate the “black box” determining observable behavior.”. But upon rereading the paper I found that this was a little underdeveloped in the manuscript. I feel that the authors can do more to draw out this contribution further. Unless I’m mistaken, I think this is aligned, or at least overlaps, with the reviewer’s point number 4 (see section c), and even with point 5 given that, to me, it seems the different affordances in the various games and their variants are what can be used in combination to illuminate the “black box” of behaviour in economic games.

Apart from these, please also address the reviewer’s other concerns, many of which I agree with but which I do not repeat here.

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.

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,

Farid Anvari

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.

The main aim of this paper is to provide a terminological framework for discussing and analyzing economic games. The authors provide this framework as a tool for optimizing research on topics which economic games are typically used to investigate. The title and conclusion focus on prosocial behavior. But the discussion is much broader, and prosociality is not the only focus in the body of the work.

This paper is difficult to evaluate. The reason for this is that the intended audience is a little unclear. On the one hand, the paper presents an overview of concepts from game theory and interdependence theory, and attempts an integration of these concepts, such that one might expect the paper to be intended for theoreticians. However, the presentation is at a very high level of abstraction; the integration is incomplete; and the presentation does not include any of the formal tools that many theorists in the area would typically use. On the other hand, the paper presents an introductory overview of basic concepts used to analyse games (including equilibria, utility, and rational choice), such that one might expect the paper to be intended for experimental social psychologists. However, the presentation of these basic concepts in section 1 assumes a lot of knowledge about different games (e.g. prisoners dilemma, dictators game) that a naïve experimentalist wouldn’t have. Thus, I feel the paper would be very much improved by focusing on one of these two audiences.

In addition, I have the following major and minor concerns

Major concerns:

1. There is a major conceptual confusion about how one should interpret the results of experiments using economic games. In the introduction, the authors argue that the value of economic games in experimental work is that behavioral results can be interpreted as direct measures of psychological states/motives. They say: “economic games are objective in terms of the degree to which observations can be directly interpreted. For instance, in the Dictator Game – one of the most commonly used games (see Section II.1 for details) – a dictator can split an endowment between herself and a recipient who has no veto power and must thus accept the dictator’s split. By implication, a dictator giving nothing is acting selfishly in the most literal sense, a dictator giving exactly half is acting in a fair manner, and a dictator giving everything is acting altruistically. Thus, in an almost behaviorist fashion, the observable behavior itself (i.e., the amount the dictator gives) serves as the variable of interest that can be directly interpreted.” In the conclusion, precisely the opposite is asserted: “Closely related, different psychological processes (including social motives) often lead to the same behavior. Thus, from observing a player’s behavior in a single game, it is rarely possible to conclusively determine a particular underlying process.” These quotes are examples, but I believe that this contradiction runs throughout the work, and the meaning of behavioral results for theory is never really clear. This is an extremely important contradiction to resolve, given the theoretical ambitions of the work. For what it’s worth, I believe the quote from the conclusion is the correct one. Results from economics games can’t be interpreted as directly revealing psychological states.
2. There is some important confusion about the terms “payoff”, “outcome”, “utility”, and the relationships between them. As I understand decision theory, payoff and outcome are generally synonymous in choices with monetary outcomes, and both refer to the numeric value the player receives. Utility refers to the level of subjective satisfaction, or dissatisfaction, the player attributes to the outcome. Thus, the term “utility function” refers to the mathematical function that transforms numeric outcomes into subjective levels of satisfaction. This is the standard interpretation of these terms in classical utility theory and prospect theory, for example. This does not appear to be the authors understanding of these terms, and the authors use payoffs and utility interchangeably. See for example note 3. This seems non-standard, at least from the perspective of decision theory. I think this is important, because it introduces some confusion into how the authors describe concepts from game theory. For instance, when discussing the notion of rationality, the authors say “Game Theory does not dictate that outcomes must be ordered according to their material consequences”, yet it is not certain what this means, if there can be no separate concept of utility representing how outcomes are subjectively evaluated. Put simply, if utilities = payoffs = outcomes, then how else could outcomes be ordered, if not in terms of material consequences? The paper should more clearly define what is meant by these terms.
3. The structure of the paper impedes readability. The earlier presentation of the theoretical content relies heavily on examples of games, which are not fully described until the later section. Thus, the examples do not do the work to illuminate the concepts that they otherwise could, and more naïve readers may not find the examples useful, prior to their complete description in the later sections. It may be worth considering re-ordering these sections.
4. It is unclear what work is to be done with the conceptual framework provided in the paper. One way theories can be useful is by providing precise formalizations for prediction and explanation in a domain. Indeed this is what classical game theory does, as a field of applied mathematics. Formalization is not the aim of the present paper. Nor is formalization strictly necessary, for theory to be useful. Informal theory can be very useful in three ways. a. Informal theory may be used to simplify phenomena into easily distinguishable categories, so we can identify underlying unity where none was apparent before. This paper does not provide this. Rather, it introduces a terminological system that increases the number of terms we can use to describe what is already describable within a simpler system, and there is substantial overlap between different terms in that system. b. Informal theory can also be useful if it can carve out new phenomena, or be used to identify things we could not identify without the theory. The terminological framework provided here does not clearly offer this. Most of the terms presented here map onto already existing folk and social psychological ideas about social interactions. c. Informal theory can also explain important sources of heterogeneity, which were previously unexplained. This is where I see the greatest potential for the presented framework. However, I feel this purpose is not pursued enough in the current manuscript. Wherever heterogeneity comes up, the argument very quickly moves on, and we do not clearly see how the terminology developed in the paper could be used to spell out the psychological variation underlying individual differences in behavior in economic games, or how the framework could be used to develop experimental predictions in this area. The overall result is that the terminology comes across as inert.
5. The application of the concept of affordances appears promising for understanding social motives etc. in the context of economic games. However, the concept is underdeveloped here. In perception research, an affordance describes how the environment places informational constraints on perceptual experience and behavior. When alluding to a possible theoretical role for affordances, the authors say “Future research is needed to systematically investigate individuals’ subjective perceptions of the affordances in different games, and how these perceptions ultimately translate to behavior”. However, as a set of ecological/informational constraints, an affordance is not something that can be subjectively perceived (or processed). Put another way, affordances determine (or constitute) the kinds of experiences one has, and the kinds of actions which are elicited in a context, and are not themselves the content of experience or cognition. They are the products of the relationship between the environment and the kind of agent/organism acting within it. Translated to the present context, games could indeed be understood as sets of affordances for social behavior. This is an idea with a lot of potential, but the idea would need to be expressed more consistently and more clearly than in the present manuscript. And importantly, this idea would need to be more explicitly related to the terminological framework presented in the body of the work.

Minor concerns:

1. The paper claims in a number of sections that it provides the first theoretical overview of game scenarios. However, there are many book and paper length treatments of this topic. It is also stated that the paper provides the first theoretical introduction to economic games. This is an odd statement, since most of the games described here were developed within applied mathematics, which is purely theoretical, and were not initially developed as experimental paradigms.
2. There is a conceptual confusion about how to interpret behavior within classic game theory. In the introduction the authors present game theory in terms of mental concepts and events. For example, they say: “Game Theory provides a formal analysis of games and behavior in games that takes into account all players’ preferences (i.e., their attitudes towards the distribution of payoffs between themselves and others)”. However, in game theory, games are mathematical objects, and equilibria etc are defined using mathematical proofs, so the allusion to psychological states here seems inappropriate.
3. The figure showing the number of papers published referring to economic games needs to be contextualized and normalized somehow. There are more papers published generally in all fields in this same time, and this might be all that the figure shows.
4. The term strategy is used inconsistently. In some places it refers to a specific action in a game. In others, the range of possible actions.
5. The concept of rationality is inconsistent throughout the paper. In some places, it is interpreted to mean something like utility maximization. In others, it is explicitly defined in terms of game theoretic equilibria. And the authors are inconsistent in how much they see these two concepts overlapping.
6. The title could be improved. There are no formal models provided in the paper, and the focus is much broader than prosociality.

My main conclusion is that I think this paper could be valuable, if it could provide a way to cash out the verbal theoretical work it describes. It could do this with a more detailed discussion of how the theory helps to explain individual differences in performance in games; how good knowledge of individual psychological processes can be used to make inferences in different situations in which games are used; or how this work could inform formal theories. The promising idea of analyzing games in terms of affordances could also be pursued in more detail, and more consistently linked to the rest of the theoretical content. Basically, the terminological framework itself needs to be put to more work, so that readers can see more clearly how they could use it.

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**

Dec 11, 2020

Dear Dr. Anvari,

Thank you very much for the positive assessment of our manuscript “Economic games: Modeling social interactions and understanding prosocial behavior” (previously entitled “Economic games: Modeling, measuring, and understanding prosocial behavior”; Manuscript ID: 1187720) and the invitation to submit a revision. We are very grateful for your and the reviewer’s constructive comments and suggestions. Gladly, we were able to address all of these to further strengthen the manuscript. In the “revision notes” appended below, we have summarized how we have dealt with each of the specific points raised and suggestions made.

We hope to have responded to your and the reviewers’ comments in a satisfactory manner. We are grateful for the opportunity to better our work.

Sincerely,

Isabel Thielmann, Robert Böhm, Marion Ott, and Ben Hilbig

**Revision Notes**

**Editor’s comments:**

I have now received the review of your manuscript, “Economic games: Modeling, measuring, and understanding prosocial behavior” from a qualified researcher. I also independently read the manuscript before consulting these review and again afterwards. I agree that your manuscript has potential to make a good contribution 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.

\*\* *Response*: We thank you very much for the positive assessment of our manuscript and the opportunity to submit a revision. In what follows, we summarize how we have dealt with each specific point raised by you and the reviewer. Moreover, we have highlighted all changes made in the manuscript using track-change mode. The page numbers provided below refer to the manuscript including tracked changes.

1. The reviewer notes that the intended audience isn’t clear. Is the paper aimed at theoreticians or naïve experimentalist social psychologists? After having read the manuscript twice, once before sending out for review and once after receiving the review, my impression is that the paper is aimed at the latter (i.e., naïve experimental psychologists). Of course, I may be wrong. However, if this is the case then there’s a lot of assumed knowledge. I recommend that the authors make clear the intended audience and the necessary adjustments to the manuscript (as per the reviewer’s recommendations).

*\*\* Response*: We thank you and the reviewer for pointing us to this important issue. Indeed, the review is sought to provide an introduction to the use of games for naïve experimentalists in psychology and beyond. We now make this explicit at several places in the manuscript (see, e.g., p. 6 and p. 9). Moreover, we have paid great attention to avoid assuming any knowledge that readers may not (yet) have in the early paragraphs of the manuscript.

1. When I first read the paper my thought was that the biggest contribution of the paper is best expressed by the following quote from the concluding remarks: “from observing a player’s behavior in a single game, it is rarely possible to conclusively determine a particular underlying process. However, a specific elegance of games is that one can combine different games or game variants to isolate certain processes and social motives […] Overall, combining different games and/or game variants – and comparing corresponding behaviors – will often provide more fine-grained insights into the psychological processes underlying players’ choices and thereby help illuminate the “black box” determining observable behavior.”. But upon rereading the paper I found that this was a little underdeveloped in the manuscript. I feel that the authors can do more to draw out this contribution further. Unless I’m mistaken, I think this is aligned, or at least overlaps, with the reviewer’s point number 4 (see section c), and even with point 5 given that, to me, it seems the different affordances in the various games and their variants are what can be used in combination to illuminate the “black box” of behaviour in economic games.

*\*\* Response*: We thank you for this valuable suggestion and agree that this contribution should be developed further in the manuscript. Thus, we now emphasize at several places in the manuscript that the observable behavior in a single game may be produced by different psychological processes (including social motives) and that combining one game with another game (or variant thereof) may help to solve this issue (e.g., p. 25). Moreover, in our Section II.3 (“Concluding remarks”; pp. 52), we now more clearly spell out that an affordance-based perspective can help isolating certain psychological processes (thus illuminating the “black box” underlying behavior) and we also provide some additional examples from prior research that combined games to achieve this goal.

**Reviewer 1**

The main aim of this paper is to provide a terminological framework for discussing and analyzing economic games. The authors provide this framework as a tool for optimizing research on topics which economic games are typically used to investigate. The title and conclusion focus on prosocial behavior. But the discussion is much broader, and prosociality is not the only focus in the body of the work.

This paper is difficult to evaluate. The reason for this is that the intended audience is a little unclear. On the one hand, the paper presents an overview of concepts from game theory and interdependence theory, and attempts an integration of these concepts, such that one might expect the paper to be intended for theoreticians. However, the presentation is at a very high level of abstraction; the integration is incomplete; and the presentation does not include any of the formal tools that many theorists in the area would typically use. On the other hand, the paper presents an introductory overview of basic concepts used to analyse games (including equilibria, utility, and rational choice), such that one might expect the paper to be intended for experimental social psychologists. However, the presentation of these basic concepts in section 1 assumes a lot of knowledge about different games (e.g. prisoners dilemma, dictators game) that a naïve experimentalist wouldn’t have. Thus, I feel the paper would be very much improved by focusing on one of these two audiences.

*\*\* Response*: We thank the reviewer for pointing us to this issue. As detailed in our response to the Editor’s comment #1, we now make explicit that the paper is aimed at providing an introduction to economic games for researchers in psychology who may want to apply economic games in their empirical studies (e.g., pp. 6 and 9). Of course, for empirical researchers too, understanding the theoretical underpinnings of games is crucial, which is why we give an introduction to both the theoretical aspects of games and their practical application. However, we understand that especially in the Introduction, our presentation of the different games may have assumed some (prior) knowledge by readers. In revising the manuscript, we have paid great attention to explain all games/concepts when referring to them first.

In addition, I have the following major and minor concerns

Major concerns:

1. There is a major conceptual confusion about how one should interpret the results of experiments using economic games. In the introduction, the authors argue that the value of economic games in experimental work is that behavioral results can be interpreted as direct measures of psychological states/motives. They say: “economic games are objective in terms of the degree to which observations can be directly interpreted. For instance, in the Dictator Game – one of the most commonly used games (see Section II.1 for details) – a dictator can split an endowment between herself and a recipient who has no veto power and must thus accept the dictator’s split. By implication, a dictator giving nothing is acting selfishly in the most literal sense, a dictator giving exactly half is acting in a fair manner, and a dictator giving everything is acting altruistically. Thus, in an almost behaviorist fashion, the observable behavior itself (i.e., the amount the dictator gives) serves as the variable of interest that can be directly interpreted.” In the conclusion, precisely the opposite is asserted: “Closely related, different psychological processes (including social motives) often lead to the same behavior. Thus, from observing a player’s behavior in a single game, it is rarely possible to conclusively determine a particular underlying process.” These quotes are examples, but I believe that this contradiction runs throughout the work, and the meaning of behavioral results for theory is never really clear. This is an extremely important contradiction to resolve, given the theoretical ambitions of the work. For what it’s worth, I believe the quote from the conclusion is the correct one. Results from economics games can’t be interpreted as directly revealing psychological states.

*\*\* Response:* We thank the reviewer for hinting us at this inconsistency. In the revised version, we no longer imply that behavior in games can be directly interpreted but only that it can be directly observed in an almost behaviorist fashion (pp. 6/7). Moreover, we have clarified throughout that definite conclusions about the underlying processes and motives driving decisions in games may only be obtained if multiple games are studied in combination (see also our response to the Editor’s comment #2).

1. There is some important confusion about the terms “payoff”, “outcome”, “utility”, and the relationships between them. As I understand decision theory, payoff and outcome are generally synonymous in choices with monetary outcomes, and both refer to the numeric value the player receives. Utility refers to the level of subjective satisfaction, or dissatisfaction, the player attributes to the outcome. Thus, the term “utility function” refers to the mathematical function that transforms numeric outcomes into subjective levels of satisfaction. This is the standard interpretation of these terms in classical utility theory and prospect theory, for example. This does not appear to be the authors understanding of these terms, and the authors use payoffs and utility interchangeably. See for example note 3. This seems non-standard, at least from the perspective of decision theory. I think this is important, because it introduces some confusion into how the authors describe concepts from game theory. For instance, when discussing the notion of rationality, the authors say “Game Theory does not dictate that outcomes must be ordered according to their material consequences”, yet it is not certain what this means, if there can be no separate concept of utility representing how outcomes are subjectively evaluated. Put simply, if utilities = payoffs = outcomes, then how else could outcomes be ordered, if not in terms of material consequences? The paper should more clearly define what is meant by these terms.

*\*\* Response:* We thank the reviewer for identifying this potential source of confusion. In fact, we did not mean to imply that outcomes are equal to payoffs and utilities, respectively. Rather, whereas outcomes refer to material (e.g., monetary) or non-material consequences of players’ behaviors, payoffs and utilities, respectively, refer to a player’s numerical representation of their preferences. Stated differently, we use “payoff = utility”, and distinguish this from “outcome = consequence” (e.g., money at stake). As such, our usage of these terms is in line with their common usage in Game Theory, which we have clarified (by also referring to a corresponding reference) in our Footnote 2 (former Footnote 3) on p. 10. Moreover, at several places in the main text, we now explicitly refer to the relation between payoff, utility, and outcomes as used in different frameworks (i.e., Game Theory and Interdependence Theory; see, e.g., pp. 10 and 26) and we have also replaced the term payoff by the term outcome in Section II whenever we do not specifically refer to a game-theoretic analysis (given that the term “outcome” may be less ambiguous, but still correct given that payoffs are usually represented by monetary or non-monetary outcomes in experimental research; see p. 26). We hope these changes will prevent misunderstandings by readers.

1. The structure of the paper impedes readability. The earlier presentation of the theoretical content relies heavily on examples of games, which are not fully described until the later section. Thus, the examples do not do the work to illuminate the concepts that they otherwise could, and more naïve readers may not find the examples useful, prior to their complete description in the later sections. It may be worth considering re-ordering these sections.

*\*\* Response:* We thank the reviewer for hinting us at this issue. As also described in our response to the reviewer’s comment #1, we have paid great attention to describe the games early on in the Introduction when referring to them first. As such, we decided to keep the general structure of the paper as is, given that the detailed introduction into games (i.e., Section II, starting on p. 25) rests on the (theoretical) concepts introduced earlier. Thus, it was not possible to simply re-ordering the paper.

1. It is unclear what work is to be done with the conceptual framework provided in the paper. One way theories can be useful is by providing precise formalizations for prediction and explanation in a domain. Indeed this is what classical game theory does, as a field of applied mathematics. Formalization is not the aim of the present paper. Nor is formalization strictly necessary, for theory to be useful. Informal theory can be very useful in three ways. a. Informal theory may be used to simplify phenomena into easily distinguishable categories, so we can identify underlying unity where none was apparent before. This paper does not provide this. Rather, it introduces a terminological system that increases the number of terms we can use to describe what is already describable within a simpler system, and there is substantial overlap between different terms in that system. b. Informal theory can also be useful if it can carve out new phenomena, or be used to identify things we could not identify without the theory. The terminological framework provided here does not clearly offer this. Most of the terms presented here map onto already existing folk and social psychological ideas about social interactions. c. Informal theory can also explain important sources of heterogeneity, which were previously unexplained. This is where I see the greatest potential for the presented framework. However, I feel this purpose is not pursued enough in the current manuscript. Wherever heterogeneity comes up, the argument very quickly moves on, and we do not clearly see how the terminology developed in the paper could be used to spell out the psychological variation underlying individual differences in behavior in economic games, or how the framework could be used to develop experimental predictions in this area. The overall result is that the terminology comes across as inert.

*\*\* Response:* We understand that our presentation was not clear in highlighting the advantages of viewing games through the lens of affordances. To overcome this issue, we now mention these advantages more explicitly in several places in the manuscript (e.g., pp. 25 and 52). Moreover, in delineating how affordances in games can be used to determine which games should be combined to isolate certain psychological processes and motives (see our response to the Editor’s comment #2 and the reviewer’s comment #5 below), the benefits of our theoretical perspective are carved out automatically.

1. The application of the concept of affordances appears promising for understanding social motives etc. in the context of economic games. However, the concept is underdeveloped here. In perception research, an affordance describes how the environment places informational constraints on perceptual experience and behavior. When alluding to a possible theoretical role for affordances, the authors say “Future research is needed to systematically investigate individuals’ subjective perceptions of the affordances in different games, and how these perceptions ultimately translate to behavior”. However, as a set of ecological/informational constraints, an affordance is not something that can be subjectively perceived (or processed). Put another way, affordances determine (or constitute) the kinds of experiences one has, and the kinds of actions which are elicited in a context, and are not themselves the content of experience or cognition. They are the products of the relationship between the environment and the kind of agent/organism acting within it. Translated to the present context, games could indeed be understood as sets of affordances for social behavior. This is an idea with a lot of potential, but the idea would need to be expressed more consistently and more clearly than in the present manuscript. And importantly, this idea would need to be more explicitly related to the terminological framework presented in the body of the work.

*\*\* Response:* We thank the reviewer for encouraging us to carve out the advantages of taking an affordance-based perspective on economic games. To this end, we now make explicit that games can be understood as providing affordances for behavior and that considering games through the lens of affordances allows deriving testable predictions about which person factors (e.g., personality traits, motives) may become expressed in behavior in a certain game (see, e.g., our introductory paragraphs in Section II; p. 25). Moreover, in the description of the games in Section II, we more thoroughly detail which affordances are involved in a game and why. In this regard, we also refer to empirical evidence that allows drawing conclusions on the usefulness of the theoretically-derived affordances for understanding game behavior in our review of typical research findings per game. Finally, we have added a paragraph in our section II.3 “Concluding remarks” (p. 52) in which we clarify the advantages of considering games in terms of the affordances they provide.

Concerning the reviewer’s comment that affordances are entities of the situation that are not subject to individuals’ perceptions, we have considerably revised the corresponding paragraph on p. 52 to clarify that different perceptions of the features of a game may lead certain affordances to become more or less prominent in driving behavior.

Minor concerns:

1. The paper claims in a number of sections that it provides the first theoretical overview of game scenarios. However, there are many book and paper length treatments of this topic. It is also stated that the paper provides the first theoretical introduction to economic games. This is an odd statement, since most of the games described here were developed within applied mathematics, which is purely theoretical, and were not initially developed as experimental paradigms.

*\*\* Response:* Although we maintain that there is no paper, chapter, or book available that provides a theoretically-founded and practical introduction into games that is specifically directed at experimental psychologists, we no longer imply by our wording that we are the first to provide such a summary.

1. There is a conceptual confusion about how to interpret behavior within classic game theory. In the introduction the authors present game theory in terms of mental concepts and events. For example, they say: “Game Theory provides a formal analysis of games and behavior in games that takes into account all players’ preferences (i.e., their attitudes towards the distribution of payoffs between themselves and others)”. However, in game theory, games are mathematical objects, and equilibria etc are defined using mathematical proofs, so the allusion to psychological states here seems inappropriate.

*\*\* Response:* We apologize for the ambiguous wording. To address this issue, we have considerably revised the respective paragraph on p. 7, now strictly sticking with game-theoretical terms without alluding to attitudes or other psychological concepts.

1. The figure showing the number of papers published referring to economic games needs to be contextualized and normalized somehow. There are more papers published generally in all fields in this same time, and this might be all that the figure shows.

*\*\* Response:* We agree that the figure might not be very conclusive in itself and have thus removed it.

1. The term strategy is used inconsistently. In some places it refers to a specific action in a game. In others, the range of possible actions.

*\*\* Response:* We thank the reviewer for hinting us at this inconsistency. We now use the term strategy consistently throughout the paper, referring to a complete plan of action for the game. In turn, to refer to a player’s available strategies, we refer to “strategy set” or “set of strategies”.

1. The concept of rationality is inconsistent throughout the paper. In some places, it is interpreted to mean something like utility maximization. In others, it is explicitly defined in terms of game theoretic equilibria. And the authors are inconsistent in how much they see these two concepts overlapping.

*\*\* Response:* We understand that our use of the word “rational” may cause some confusion. To address this issue, we have deleted the sentence that connects rational choices with equilibrium solutions (pp. 14/15). Moreover, we replaced the expression “selfish-rational” with the term “selfish”.

1. The title could be improved. There are no formal models provided in the paper, and the focus is much broader than prosociality.

*\*\* Response:* As suggested, we have changed our title to “Economic games: Modeling social interactions and understanding prosocial behavior”. We hope that, this way, the title better conveys that games are sought to specifically model different kinds of social situations, but not to formally model prosocial behavior.

My main conclusion is that I think this paper could be valuable, if it could provide a way to cash out the verbal theoretical work it describes. It could do this with a more detailed discussion of how the theory helps to explain individual differences in performance in games; how good knowledge of individual psychological processes can be used to make inferences in different situations in which games are used; or how this work could inform formal theories. The promising idea of analyzing games in terms of affordances could also be pursued in more detail, and more consistently linked to the rest of the theoretical content. Basically, the terminological framework itself needs to be put to more work, so that readers can see more clearly how they could use it.

*\*\* Response:* We thank the reviewer for these helpful suggestions. As detailed in our responses to the reviewer’s comments above, we have considerably revised the paper to more clearly spell out the advantages of the theoretical perspective proposed in the paper and the implications a corresponding experimental approach can have for the understanding of the psychological processes involved in prosocial behavior as measured in games.

**Editor second decision—Revise & Resubmit**

Jan 7, 2021

Dear Dr. Thielmann,

I have now received another round of review of your manuscript, “Economic games: Modeling social interactions and understanding prosocial behavior” from the same reviewer. I would like to thank this reviewer for their timely evaluation and for their expertise and time in helping me to evaluate the manuscript. I also independently read the manuscript before consulting the review. Once again, I agree that your manuscript has important strengths and, although most of the original concerns have been satisfactorily addressed, there are still some issues that remain but which I think can be relatively easily overcome. I therefore encourage you to submit a revised version for further consideration at Collabra: Psychology.

The reviewer did an outstanding job in their reviews. I will highlight issues that were particularly salient for me without reiterating the reviewer’s concerns, so please pay attention to the reviewer’s points as well. In your resubmission, please include a document with a point-by-point response to both the points I list here and the reviewer’s comments, outlining each change made in your manuscript or providing a suitable rebuttal–ideally with a track-changes version of the manuscript if possible. Note that the reviewer’s remaining concerns are presented in response to your responses to the previous concerns. My remaining concerns are as follow:

As per the reviewer’s response to point 1, I would also like to argue that the authors’ claims about the benefits of games is too strongly stated given what the paper later argues for. For example, the paper states that “economic games are objective in terms of the degree to which the dependent variable is directly observable. […] the observable behavior itself (i.e., the amount the dictator gives to the recipient) serves as the variable of interest that can be directly observed. As a consequence – and advantage –, there is little to no need for wide-ranging additional presumptions or theory-dependent assignments of constructs to observations” (p. 5). However, earlier, the paper stated that the key reason for the advantages of economic games not being fully exploited in psychological research is that “[…] theoretically-founded guidelines for the use of games in empirical research are missing, as is a broad overview of […] (iii) which psychological processes they afford to become expressed in behavior and, thus, what the games actually measure” (p. 5).

Most importantly, a main contribution of the paper is to provide a way for psychological researchers to use the games in various combinations in order to draw conclusions about psychological states. That is, the paper stipulates the different affordances of the games and thus the psychological states/processes that the games can be inferred to measure, particularly when used in certain combinations. This requires assumptions about how the behaviours observed in the games represent those psychological states. Psychological states (or constructs) are not observable; rather, we use observables (whether they be self-reports of subjective experiences or behaviours in economic games) to draw inferences about the unobservable psychological constructs in which we are interested and this always involves assumptions, unless one takes a strict behaviorist stance and commits to drawing inferences only about the behaviours observed. The authors do not seem to take such a behaviorist stance, nor should they in my opinion. For example, the paper states that “Thus, as detailed below (Section I.3), different (classes of) games provide different affordances for behavior, and knowing the type and structure of a game implies which psychological processes are afforded and may ultimately become expressed in behavior” (p. 10). Similarly, the framework provides a way for researchers to draw inferences about people’s motives, goals, values, and preferences, based on their behaviour in the games (pp. 17-18).

All of this is to say that what the authors propose goes beyond the advantage of the games presented earlier in the paper by going beyond what is directly observable to making claims about psychological states/processes that are not directly observable but inferred from observed behaviour (i.e., which “may ultimately become expressed in behavior”). For these reasons, I agree with the reviewer (see their response to the authors’ response to point 1).

I believe that the Table 2 shows the affordances provided by each game and its variants, but I recommend that the authors also point out the affordances for each game and its variants in the main text. This is done for many of the games already (e.g., trust game, prisoner’s dilemma), but some (e.g., game variants of ultimatum game) don’t include a discussion of the affordances provided. Likewise, I suggest that, where possible, to include some descriptions of how typical findings for the games reflect the different affordances the game provides (e.g., as done for typical findings in the dictator game and trust game).

Minor points.

The paper states that, “As such, we aim to foster high-quality, theory-based, and replicable research on the many aspects of human prosociality” (p. 7). But I cannot see how the proposed framework would improve replicability.

The paper states that, “Besides describing the structural features and theoretical underpinnings of the games, we derive the (sub-)affordances each game involves. This provides testable predictions about which psychological processes (including social motives) should be afforded in a game to guide behavior […]” (p. 23). But I don’t see how this is testable? How can researchers test predictions about which psychological processes are afforded in a game? Perhaps by using self-reports of, for example, personality traits?

p. 40: “Taking, in turn, may be an expression of unconditional concern for others through the motives of greed, competitiveness, and/or spite”. This sentence seems to involve a contradiction: how can “taking” be an expression of unconditional concern for others through the motives of greed etc? Isn’t taking an expression of unconcern for others? Perhaps I’m misunderstanding something here.

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.

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,

Farid Anvari

**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.

\*\*\*Reviewer: I thank the authors for so systematically addressing my first round of comments. I think the paper is much improved. The intended audience is clearer, terms are more clearly explained for more naïve readers, and the paper is theoretically sharper. I do however have a few follow-ups, which I hope the authors don’t mind addressing. These concern definition and use of the concepts of outcome and utility, plus a few more minor points.

The main aim of this paper is to provide a terminological framework for discussing and analyzing economic games. The authors provide this framework as a tool for optimizing research on topics which economic games are typically used to investigate. The title and conclusion focus on prosocial behavior. But the discussion is much broader, and prosociality is not the only focus in the body of the work.

This paper is difficult to evaluate. The reason for this is that the intended audience is a little unclear. On the one hand, the paper presents an overview of concepts from game theory and interdependence theory, and attempts an integration of these concepts, such that one might expect the paper to be intended for theoreticians. However, the presentation is at a very high level of abstraction; the integration is incomplete; and the presentation does not include any of the formal tools that many theorists in the area would typically use. On the other hand, the paper presents an introductory overview of basic concepts used to analyse games (including equilibria, utility, and rational choice), such that one might expect the paper to be intended for experimental social psychologists. However, the presentation of these basic concepts in section 1 assumes a lot of knowledge about different games (e.g. prisoners dilemma, dictators game) that a naïve experimentalist wouldn’t have. Thus, I feel the paper would be very much improved by focusing on one of these two audiences. \*\* Response: We thank the reviewer for pointing us to this issue. As detailed in our response to the Editor’s comment #1, we now make explicit that the paper is aimed at providing an introduction to economic games for researchers in psychology who may want to apply economic games in their empirical studies (e.g., pp. 6 and 9). Of course, for empirical researchers too, understanding the theoretical underpinnings of games is crucial, which is why we give an introduction to both the theoretical aspects of games and their practical application. However, we understand that especially in the Introduction, our presentation of the different games may have assumed some (prior) knowledge by readers. In revising the manuscript, we have paid great attention to explain all games/concepts when referring to them first.

\*\*\*Reviewer: The audience is clear and the paper is more accessible.

In addition, I have the following major and minor concerns Major concerns:

1. There is a major conceptual confusion about how one should interpret the results of experiments using economic games. In the introduction, the authors argue that the value of economic games in experimental work is that behavioral results can be interpreted as direct measures of psychological states/motives. They say: “economic games are objective in terms of the degree to which observations can be directly interpreted. For instance, in the Dictator Game – one of the most commonly used games (see Section II.1 for details) – a dictator can split an endowment between herself and a recipient who has no veto power and must thus accept the dictator’s split. By implication, a dictator giving nothing is acting selfishly in the most literal sense, a dictator giving exactly half is acting in a fair manner, and a dictator giving everything is acting altruistically. Thus, in an almost behaviorist fashion, the observable behavior itself (i.e., the amount the dictator gives) serves as the variable of interest that can be directly interpreted.” In the conclusion, precisely the opposite is asserted: “Closely related, different psychological processes (including social motives) often lead to the same behavior. Thus, from observing a player’s behavior in a single game, it is rarely possible to conclusively determine a particular underlying process.” These quotes are examples, but I believe that this contradiction runs throughout the work, and the meaning of behavioral results for theory is never really clear. This is an extremely important contradiction to resolve, given the theoretical ambitions of the work. For what it’s worth, I believe the quote from the conclusion is the correct one. Results from economics games can’t be interpreted as directly revealing psychological states.

\*\* Response: We thank the reviewer for hinting us at this inconsistency. In the revised version, we no longer imply that behavior in games can be directly interpreted but only that it can be directly observed in an almost behaviorist fashion (pp. 6/7). Moreover, we have clarified throughout that definite conclusions about the underlying processes and motives driving decisions in games may only be obtained if multiple games are studied in combination (see also our response to the Editor’s comment #2).

\*\*\*Reviewer: This is much improved, but I still want to push back on this point. I still think the broad claim that “economic games are objective in terms of the degree to which the dependent variable is directly observable” is too strong. Perhaps we can agree that this is a matter of degree? The dictator game is one example where the behavior itself can be interpreted as an operationalization of a psychological concept - i.e. selfishness. But this isn’t true of many other games. For example, in an iterated prisoner’s dilemma game, defection could equally reflect a selfish move, or could be an instance of conditional cooperation, which is other regarding (if not prosocial). (As an aside maybe dependent variable is slightly misleading here? The dependent variable is by definition observable. But, I still sense that what you are trying to say here is that for economic games, there is a one-to-one mapping between the dependent variable, and some psychological construct like selfishness. I say this because presumably, when you interpret a dependent variable in an operationalization, you do so in terms of a psychological state.) To express my worry another way, I think the direct operationalization of (e.g.) selfishness is only (potentially) the case for a limited number of (nonetheless interesting) games, (i.e. single move dictator and ultimatum games). Since this paper also advocates for the use of other games – e.g. the prisoners dilemma and the public goods game – I still worry that this strength is positioned too generally. And in terms of the broader argument in the rest of the paper – i.e. that interdependence theory is necessary to connect the given and effective situations – maybe the case for using games in experiments doesn’t need to be stated so strongly anyway?

1. There is some important confusion about the terms “payoff”, “outcome”, “utility”, and the relationships between them. As I understand decision theory, payoff and outcome are generally synonymous in choices with monetary outcomes, and both refer to the numeric value the player receives. Utility refers to the level of subjective satisfaction, or dissatisfaction, the player attributes to the outcome. Thus, the term “utility function” refers to the mathematical function that transforms numeric outcomes into subjective levels of satisfaction. This is the standard interpretation of these terms in classical utility theory and prospect theory, for example. This does not appear to be the authors understanding of these terms, and the authors use payoffs and utility interchangeably. See for example note 3. This seems non-standard, at least from the perspective of decision theory. I think this is important, because it introduces some confusion into how the authors describe concepts from game theory. For instance, when discussing the notion of rationality, the authors say “Game Theory does not dictate that outcomes must be ordered according to their material consequences”, yet it is not certain what this means, if there can be no separate concept of utility representing how outcomes are subjectively evaluated. Put simply, if utilities = payoffs = outcomes, then how else could outcomes be ordered, if not in terms of material consequences? The paper should more clearly define what is meant by these terms.

\*\* Response: We thank the reviewer for identifying this potential source of confusion. In fact, we did not mean to imply that outcomes are equal to payoffs and utilities, respectively. Rather, whereas outcomes refer to material (e.g., monetary) or non-material consequences of players’ behaviors, payoffs and utilities, respectively, refer to a player’s numerical representation of their preferences. Stated differently, we use “payoff = utility”, and distinguish this from “outcome = consequence” (e.g., money at stake). As such, our usage of these terms is in line with their common usage in Game Theory, which we have clarified (by also referring to a corresponding reference) in our Footnote 2 (former Footnote 3) on p. 10. Moreover, at several places in the main text, we now explicitly refer to the relation between payoff, utility, and outcomes as used in different frameworks (i.e., Game Theory and Interdependence Theory; see, e.g., pp. 10 and 26) and we have also replaced the term payoff by the term outcome in Section II whenever we do not specifically refer to a game-theoretic analysis (given that the term “outcome” may be less ambiguous, but still correct given that payoffs are usually represented by monetary or non-monetary outcomes in experimental research; see p. 26). We hope these changes will prevent misunderstandings by readers.

\*\*\*Reviewer: This is much clearer. Also, the reason for my confusion is more apparent to me (reading more from a decision theory perspective). The addition of the sentence “it has been proposed that the term “preference indicator function” may better capture the meaning of utility functions..” (p 9) is particularly helpful. To better understand this issue, let me ask a clarificatory question, from which some further theoretical questions and suggestions follow. If I understand correctly, the paper asserts i) “payoff” and “utility” are interchangeable terms meaning (quote) “numerical representation of [a players] preferences” p9 ii) “outcomes” means something different, namely the concrete material value that one gets at the end of a game or round, and iii) (quote) “outcomes are understood to directly represent preferences, meaning they directly translate to payoffs”p9 From these three points, is it correctly inferred that outcomes are directly translated to payoffs, and payoffs are identical to utility? If one were to read this from a decision theoretic perspective, could a way to understand this then be to say that in Game Theory, the utility function is always assumed to be linear (with an intercept of 0 and a slope of 1)? I can see that this might be reasonable in the cases used in the paper, and in the payoff calculations presented. I can also understand that this might be a broad assumption in the Game Theoretic textbooks referred to (except Camerer 2003). So would this be a reasonable framing of the relationship between these concepts, as described in the paper? And is this a reasonable interpretation of what is meant by “preference indicator function”? Some theoretical question remain for me, then. The first concerns the way in which Interdependence Theory is presented as a compliment to game theory. On page 17 it says “whereas in interdependence theory the starting point of any consideration is the given situation, in Game Theory it is the effective situation”. From the way that the given and the effective situations are defined, I read this as saying something roughly like the following (my words): Game Theory analyzes interactive decisions in terms of payoff matrices which represent utility and not outcomes. Interdependence theory compliments Game Theory, because it provides tools to analyze how utilities are psychologically constructed from outcomes, in a way that allows us to capture social motives (and other interesting social psychological processes). From this understanding, and given the above, would it then be right to understand Interdependence Theory as a kind of expansion or psychological generalization of Game Theory, in which the assumption of a linear utility transformation is relaxed (plus some other conceptual toosl)? Would this then be a reasonable way to read the application of these concepts throughout the paper, and to understand the use of Interdependence Theory to provide psychological meat to the skeleton provided by Game Theory? This leads to my second theoretical question, which concerns the uniqueness of Interdependence Theory on this point, and whether there aren’t approaches within Game Theory (broadly understood) that aim towards the same objective? The paper explicitly references the strategy, used in Interdependence Theory, of defining social motives in terms of weightings between one’s own and other’s utility functions. This strategy is not unique to Interdependence Theory. Prominent examples include McClinktock, C. G. (1972). Social motivation: A set of propositions. Behavioral Science, 17(5), 438-455 Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. The Quarterly Journal of Economics, 117(3), 817-869 Janssenn, M. A., & Ahn, T. K. (2006).Learning, signaling, and social preferences in public goods games. Ecology and Society, 11(2). More broadly, Bayesian and Reinforcement Learning approaches to modeling choice in games also relax the assumption of a linear function relating outcomes and utilities. Examples include Masel, J. (2007). A Bayesian model of quasi-magical thinking can explain observed cooperation in the public good game. Journal of Economic Behavior and Organization, 64, 216-231; 33; Camerer, C., & Ho, T. (1999). Experience weighted attraction in normal form games. Econometrica, 67(4), 827-874; Roth, A., & Erev, I. (1995). Learning in extensive form games: experimental data and simple dynamic models in the intermediate term. Games and Economic Behavior, 8, 164-212. I think it is this work that has fueled my confusion on this point. Each of the abovementioned papers are cases in which an attempt is made to provide a psychological account of how outcomes are transformed to utilities when playing economic games; or at least the models they present include explicit utility functions for this purpose. And although they may not be canonical examples of classical Game Theory, I would have counted them within the scope of Game Theory (perhaps much) more broadly construed. (Following Camerer 2003, we might call this behavioral Game Theory). My question then is whether this work also deserves a mention, possibly as a very prominent set of approaches to psychologizing Game Theory? Mention of these approaches could be noted at the beginning of section I (p8), near where it says “In this section we will give an introduction to the two most prominent theories of games…”. At least RL theories might deserve a mention as also being as prominent (both Camerer & Ho and Roth and Erev have >2000 citations). I don’t think a full account of this literature is necessary. That would overcomplicate the narrative. But this at least explains my confusion, and may be a cause of confusion for readers who are familiar with these models, so that some reference to them is helpful. A final point on outcomes and utility is that I do perceive some remaining conceptual slips concerning whether payoff matrices represent outcomes or utilities. This is mainly when introducing the prisoner’s dilemma. For example on page 18 “To illustrate this, again consider the Prisoner’s Dilemma in Figure 1A, assuming that the payoffs in the matrix denote (monetary) outcomes.”

1. The structure of the paper impedes readability. The earlier presentation of the theoretical content relies heavily on examples of games, which are not fully described until the later section. Thus, the examples do not do the work to illuminate the concepts that they otherwise could, and more naïve readers may not find the examples useful, prior to their complete description in the later sections. It may be worth considering re-ordering these sections.

\*\* Response: We thank the reviewer for hinting us at this issue. As also described in our response to the reviewer’s comment #1, we have paid great attention to describe the games early on in the Introduction when referring to them first. As such, we decided to keep the general structure of the paper as is, given that the detailed introduction into games (i.e., Section II, starting on p. 25) rests on the (theoretical) concepts introduced earlier. Thus, it was not possible to simply re-ordering the paper.

\*\*\*Reviewer: This works much better now.

1. It is unclear what work is to be done with the conceptual framework provided in the paper. One way theories can be useful is by providing precise formalizations for prediction and explanation in a domain. Indeed this is what classical game theory does, as a field of applied mathematics. Formalization is not the aim of the present paper. Nor is formalization strictly necessary, for theory to be useful. Informal theory can be very useful in three ways. a. Informal theory may be used to simplify phenomena into easily distinguishable categories, so we can identify underlying unity where none was apparent before. This paper does not provide this. Rather, it introduces a terminological system that increases the number of terms we can use to describe what is already describable within a simpler system, and there is substantial overlap between different terms in that system. b. Informal theory can also be useful if it can carve out new phenomena, or be used to identify things we could not identify without the theory. The terminological framework provided here does not clearly offer this. Most of the terms presented here map onto already existing folk and social psychological ideas about social interactions. c. Informal theory can also explain important sources of heterogeneity, which were previously unexplained. This is where I see the greatest potential for the presented framework. However, I feel this purpose is not pursued enough in the current manuscript. Wherever heterogeneity comes up, the argument very quickly moves on, and we do not clearly see how the terminology developed in the paper could be used to spell out the psychological variation underlying individual differences in behavior in economic games, or how the framework could be used to develop experimental predictions in this area. The overall result is that the terminology comes across as inert.

\*\* Response: We understand that our presentation was not clear in highlighting the advantages of viewing games through the lens of affordances. To overcome this issue, we now mention these advantages more explicitly in several places in the manuscript (e.g., pp. 25 and 52). Moreover, in delineating how affordances in games can be used to determine which games should be combined to isolate certain psychological processes and motives (see our response to the Editor’s comment #2 and the reviewer’s comment #5 below), the benefits of our theoretical perspective are carved out automatically.

\*\*\*Reviewer: This works much better now. It is now clear how the intended reader (more naïve experimental psychologist) could use the framework to design a study using a combination of games to address their own research question. The result is more productive and informative.

1. The application of the concept of affordances appears promising for understanding social motives etc. in the context of economic games. However, the concept is underdeveloped here. In perception research, an affordance describes how the environment places informational constraints on perceptual experience and behavior. When alluding to a possible theoretical role for affordances, the authors say “Future research is needed to systematically investigate individuals’ subjective perceptions of the affordances in different games, and how these perceptions ultimately translate to behavior”. However, as a set of ecological/informational constraints, an affordance is not something that can be subjectively perceived (or processed). Put another way, affordances determine (or constitute) the kinds of experiences one has, and the kinds of actions which are elicited in a context, and are not themselves the content of experience or cognition. They are the products of the relationship between the environment and the kind of agent/organism acting within it. Translated to the present context, games could indeed be understood as sets of affordances for social behavior. This is an idea with a lot of potential, but the idea would need to be expressed more consistently and more clearly than in the present manuscript. And importantly, this idea would need to be more explicitly related to the terminological framework presented in the body of the work.

\*\* Response: We thank the reviewer for encouraging us to carve out the advantages of taking an affordance-based perspective on economic games. To this end, we now make explicit that games can be understood as providing affordances for behavior and that considering games through the lens of affordances allows deriving testable predictions about which person factors (e.g., personality traits, motives) may become expressed in behavior in a certain game (see, e.g., our introductory paragraphs in Section II; p. 25). Moreover, in the description of the games in Section II, we more thoroughly detail which affordances are involved in a game and why. In this regard, we also refer to empirical evidence that allows drawing conclusions on the usefulness of the theoretically-derived affordances for understanding game behavior in our review of typical research findings per game. Finally, we have added a paragraph in our section II.3 “Concluding remarks” (p. 52) in which we clarify the advantages of considering games in terms of the affordances they provide. Concerning the reviewer’s comment that affordances are entities of the situation that are not subject to individuals’ perceptions, we have considerably revised the corresponding paragraph on p. 52 to clarify that different perceptions of the features of a game may lead certain affordances to become more or less prominent in driving behavior.

\*\*\*Reviewer: This is clearer. The theoretical role of affordances in the paper is more prominent, and clearer. It occurs to me that the concept of affordances, as used in this paper, is a component of interdependence theory. Can this also be made clearer when the concept is introduced, and in the conclusion. I think this could pique the readers interest to learn more about this useful application of the concept.

Minor concerns:

1. The paper claims in a number of sections that it provides the first theoretical overview of game scenarios. However, there are many book and paper length treatments of this topic. It is also stated that the paper provides the first theoretical introduction to economic games. This is an odd statement, since most of the games described here were developed within applied mathematics, which is purely theoretical, and were not initially developed as experimental paradigms.

\*\* Response: Although we maintain that there is no paper, chapter, or book available that provides a theoretically-founded and practical introduction into games that is specifically directed at experimental psychologists, we no longer imply by our wording that we are the first to provide such a summary.

\*\*\*Reviewer: I would argue that this is one of the goals of Camerer 2003, especially the earlier chapters, which give a detailed and accessible introduction to the dictator and ultimatum games (the content of this presentation is not unlike the one given in the present manuscript). Camerer also works to provide the theoretical terminology (and formal tools) for other social scientists. This is arguably also the goal of e.g. Coleman 1995 – “Game Theory and It’s Applications in the Social and Behavioral Sciences”. Could the paper perhaps be situated as a contribution to this kind of introductory work, that updates (or focuses on) Interdependence Theory?

1. There is a conceptual confusion about how to interpret behavior within classic game theory. In the introduction the authors present game theory in terms of mental concepts and events. For example, they say: “Game Theory provides a formal analysis of games and behavior in games that takes into account all players’ preferences (i.e., their attitudes towards the distribution of payoffs between themselves and others)”. However, in game theory, games are mathematical objects, and equilibria etc are defined using mathematical proofs, so the allusion to psychological states here seems inappropriate.

\*\* Response: We apologize for the ambiguous wording. To address this issue, we have considerably revised the respective paragraph on p. 7, now strictly sticking with game-theoretical terms without alluding to attitudes or other psychological concepts.

\*\*\*Reviewer: I still see some tension here. For example, it is my understanding that the “effective situation” refers to something like a psychological interpretation of the game, e.g. (quote) “players are assumed to engage in a psychological transformation, which eventually results in an effective situation” on p16. In this context the paper also says (quote) “whereas in Interdependence Theory the starting point of any consideration is the given situation, in Game Theory it is the effective situation” p17. This is a more minor issue, so I don’t think anything needs to be changed, but I think it is also connected to the broader issue raised above concerning utility etc.

1. The figure showing the number of papers published referring to economic games needs to be contextualized and normalized somehow. There are more papers published generally in all fields in this same time, and this might be all that the figure shows.

\*\* Response: We agree that the figure might not be very conclusive in itself and have thus removed it. \*\*\*Reviewer: This is resolved.

1. The term strategy is used inconsistently. In some places it refers to a specific action in a game. In others, the range of possible actions.

\*\* Response: We thank the reviewer for hinting us at this inconsistency. We now use the term strategy consistently throughout the paper, referring to a complete plan of action for the game. In turn, to refer to a player’s available strategies, we refer to “strategy set” or “set of strategies”. \*\*\*Reviewer: This is resolved.

1. The concept of rationality is inconsistent throughout the paper. In some places, it is interpreted to mean something like utility maximization. In others, it is explicitly defined in terms of game theoretic equilibria. And the authors are inconsistent in how much they see these two concepts overlapping.

\*\* Response: We understand that our use of the word “rational” may cause some confusion. To address this issue, we have deleted the sentence that connects rational choices with equilibrium solutions (pp. 14/15). Moreover, we replaced the expression “selfish-rational” with the term “selfish”. \*\*\*Reviewer: This is resolved

1. The title could be improved. There are no formal models provided in the paper, and the focus is much broader than prosociality.

\*\* Response: As suggested, we have changed our title to “Economic games: Modeling social interactions and understanding prosocial behavior”. We hope that, this way, the title better conveys that games are sought to specifically model different kinds of social situations, but not to formally model prosocial behavior. \*\*\*Reviewer: There may be readers for whom “model” strongly cues formal theory, but I leave this entirely to the authors’ discretion.

My main conclusion is that I think this paper could be valuable, if it could provide a way to cash out the verbal theoretical work it describes. It could do this with a more detailed discussion of how the theory helps to explain individual differences in performance in games; how good knowledge of individual psychological processes can be used to make inferences in different situations in which games are used; or how this work could inform formal theories. The promising idea of analyzing games in terms of affordances could also be pursued in more detail, and more consistently linked to the rest of the theoretical content. Basically, the terminological framework itself needs to be put to more work, so that readers can see more clearly how they could use it.

\*\* Response: We thank the reviewer for these helpful suggestions. As detailed in our responses to the reviewer’s comments above, we have considerably revised the paper to more clearly spell out the advantages of the theoretical perspective proposed in the paper and the implications a corresponding experimental approach can have for the understanding of the psychological processes involved in prosocial behavior as measured in games.

\*\*\*Reviewer: This is more successful.

**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**

Jan 19, 2021

Dear Dr. Anvari,

Thank you very much for the positive assessment of our revised manuscript “Economic games: An introduction and guide for research” (previously entitled “Economic games: Modeling social interactions and understanding prosocial behavior”; Manuscript ID: 1187720) and the invitation to submit another revision. We are once more very grateful for your and the reviewer’s constructive comments and suggestions. Gladly, we were able to address all of these to further strengthen the manuscript. In the “revision notes” attached to our submission, we have summarized how we have dealt with each of the specific points raised and suggestions made. Moreover, we are attaching a manuscript version to our submission in which all changes made are tracked using Track Change Mode.

We hope to have responded to your and the reviewer’s comments in a satisfactory manner and that our manuscript is now ready for publication in Collabra: Psychology. We are grateful for the opportunity to better our work.

Sincerely,

Isabel Thielmann, Robert Böhm, Marion Ott, and Ben Hilbig

**Author Revision notes:**

**Revision Notes**

**Editor’s comments:**

Once again, I agree that your manuscript has important strengths and, although most of the original concerns have been satisfactorily addressed, there are still some issues that remain but which I think can be relatively easily overcome. I therefore encourage you to submit a revised version for further consideration at Collabra: Psychology.

\*\*\**Response*: We thank you for your positive feedback on our revised manuscript. We are happy that most of the original concerns have been addressed in a satisfactory manner, and we hope that the additional changes detailed below have solved all remaining issues.

1. As per the reviewer’s response to point 1, I would also like to argue that the authors’ claims about the benefits of games is too strongly stated given what the paper later argues for. For example, the paper states that “economic games are objective in terms of the degree to which the dependent variable is directly observable. […] the observable behavior itself (i.e., the amount the dictator gives to the recipient) serves as the variable of interest that can be directly observed. As a consequence – and advantage –, there is little to no need for wide-ranging additional presumptions or theory-dependent assignments of constructs to observations” (p. 5). However, earlier, the paper stated that the key reason for the advantages of economic games not being fully exploited in psychological research is that “[…] theoretically-founded guidelines for the use of games in empirical research are missing, as is a broad overview of […] (iii) which psychological processes they afford to become expressed in behavior and, thus, what the games actually measure” (p. 5).

Most importantly, a main contribution of the paper is to provide a way for psychological researchers to use the games in various combinations in order to draw conclusions about psychological states. That is, the paper stipulates the different affordances of the games and thus the psychological states/processes that the games can be inferred to measure, particularly when used in certain combinations. This requires assumptions about how the behaviours observed in the games represent those psychological states. Psychological states (or constructs) are not observable; rather, we use observables (whether they be self-reports of subjective experiences or behaviours in economic games) to draw inferences about the unobservable psychological constructs in which we are interested and this always involves assumptions, unless one takes a strict behaviorist stance and commits to drawing inferences only about the behaviours observed. The authors do not seem to take such a behaviorist stance, nor should they in my opinion. For example, the paper states that “Thus, as detailed below (Section I.3), different (classes of) games provide different affordances for behavior, and knowing the type and structure of a game implies which psychological processes are afforded and may ultimately become expressed in behavior” (p. 10). Similarly, the framework provides a way for researchers to draw inferences about people’s motives, goals, values, and preferences, based on their behaviour in the games (pp. 17-18).

All of this is to say that what the authors propose goes beyond the advantage of the games presented earlier in the paper by going beyond what is directly observable to making claims about psychological states/processes that are not directly observable but inferred from observed behaviour (i.e., which “may ultimately become expressed in behavior”). For these reasons, I agree with the reviewer (see their response to the authors’ response to point 1).

\*\*\* *Response*: We are very sorry that our exposition still triggered misunderstandings and regret that it falsely appeared inconsistent with one of the major goals of our paper, namely to show how an affordance framework can help illuminate the (otherwise unobservable) psychological processes involved in prosocial behavior. To be clear, we completely agree that the psychological processes (including social motives) producing behavior in games cannot be directly observed. However, this was not our point. What we instead intend to make clear is that the variable of interest – that is, the variable we want to explain using games, namely prosocial vs. selfish behavior – is directly observable which is a key advantage of games, especially in comparison to many other paradigms used in psychology. To clarify this, we have again considerably revised the corresponding paragraph on p. 5, which now reads as follows:

“First, the observations produced by economic games directly serve as the variable of interest and represent the to-be-measured construct – a preference – without the need for wide-ranging auxiliary assumptions. Unlike approximating a construct such as prejudice from observed differences in response times (as in the Implicit Association Test; e.g., Devine et al., 2002) or self-esteem from self-reports on questionnaire items, choices in economic games express a preference and this preference is the construct one aims to explain. Importantly, how this preference comes about – that is, the causal mechanisms producing the preference (e.g., psychological processes) – is a separate issue (though economic games also bear significant advantages in this regard, see below). In this particular respect, economic games continue the behaviorist tradition: In a classical conditioning experiment, the (strength of the) conditioned response itself directly reflects the construct of interest, namely the (strength of the) association between stimuli. The causal mechanisms producing the association are a separate issue (indeed, mechanisms involving cognition were deliberately neglected in the behaviorist tradition), but there is nonetheless an element of objectivity to the paradigm itself: Observed responses are the construct of interest.”

We hope that the repeated revision clarifies what we are trying to say. However, if you continue to think that this paragraph is misleading, we are certainly prepared to remove it from the paper.

2. I believe that the Table 2 shows the affordances provided by each game and its variants, but I recommend that the authors also point out the affordances for each game and its variants in the main text. This is done for many of the games already (e.g., trust game, prisoner’s dilemma), but some (e.g., game variants of ultimatum game) don’t include a discussion of the affordances provided. Likewise, I suggest that, where possible, to include some descriptions of how typical findings for the games reflect the different affordances the game provides (e.g., as done for typical findings in the dictator game and trust game).

\*\*\* *Response*: We thank you for this comment. As suggested, in the main text we now refer to affordances for all games and variants thereof and we also clarify for all games how typical findings can offer information on the affordances present. Moreover, we have added several references to Table 2 and our supplementary materials on the OSF in which we provide more detailed descriptions of the game variants and affordances involved.

Minor points.

The paper states that, “As such, we aim to foster high-quality, theory-based, and replicable research on the many aspects of human prosociality” (p. 7). But I cannot see how the proposed framework would improve replicability.

\*\*\* *Response*: We thank you for hinting us at this misleading statement. Originally, the replicability aspect referred to the provision of standardized instructions (on the OSF) which we hope may increase replicability of research using games. However, since this very feature of our paper is only introduced later in the text, we have removed “replicable” from the respective sentence on p. 8, which now reads: “As such, we aim to foster high-quality and theory-based research on the many aspects of human prosociality”.

The paper states that, “Besides describing the structural features and theoretical underpinnings of the games, we derive the (sub-)affordances each game involves. This provides testable predictions about which psychological processes (including social motives) should be afforded in a game to guide behavior […]” (p. 23). But I don’t see how this is testable? How can researchers test predictions about which psychological processes are afforded in a game? Perhaps by using self-reports of, for example, personality traits?

\*\*\* *Response*: We understand that such predictions may not necessarily be testable in a straightforward manner. We have thus replaced the word “testable” by “clear”. The sentence on p. 24 now reads: “This provides clear predictions about which psychological processes (including social motives) should be afforded in a game…”

p. 40: “Taking, in turn, may be an expression of unconditional concern for others through the motives of greed, competitiveness, and/or spite”. This sentence seems to involve a contradiction: how can “taking” be an expression of unconditional concern for others through the motives of greed etc? Isn’t taking an expression of unconcern for others? Perhaps I’m misunderstanding something here.

\*\*\* *Response*: We thank you for hinting us at this ambiguity. Indeed, taking should be an expression of \*low\* unconditional concern for others. We have revised the sentence correspondingly: “Taking, in turn, may be an expression of (low) unconditional concern for others through the motives of greed, competitiveness, and/or spite.” (p. 41).

**Reviewer 1:**

I thank the authors for so systematically addressing my first round of comments. I think the paper is much improved. The intended audience is clearer, terms are more clearly explained for more naïve readers, and the paper is theoretically sharper. I do however have a few follow-ups, which I hope the authors don’t mind addressing. These concern definition and use of the concepts of outcome and utility, plus a few more minor points.

\*\*\* *Response*: We thank the reviewer for the positive assessment of our revision and for the additional comments made. We are confident that addressing these has further strengthened the manuscript.

I still think the broad claim that “economic games are objective in terms of the degree to which the dependent variable is directly observable” is too strong. Perhaps we can agree that this is a matter of degree? The dictator game is one example where the behavior itself can be interpreted as an operationalization of a psychological concept - i.e. selfishness. But this isn’t true of many other games. For example, in an iterated prisoner’s dilemma game, defection could equally reflect a selfish move, or could be an instance of conditional cooperation, which is other regarding (if not prosocial). (As an aside maybe dependent variable is slightly misleading here? The dependent variable is by definition observable. But, I still sense that what you are trying to say here is that for economic games, there is a one-to-one mapping between the dependent variable, and some psychological construct like selfishness. I say this because presumably, when you interpret a dependent variable in an operationalization, you do so in terms of a psychological state.) To express my worry another way, I think the direct operationalization of (e.g.) selfishness is only (potentially) the case for a limited number of (nonetheless interesting) games, (i.e. single move dictator and ultimatum games). Since this paper also advocates for the use of other games – e.g. the prisoners dilemma and the public goods game – I still worry that this strength is positioned too generally. And in terms of the broader argument in the rest of the paper – i.e. that interdependence theory is necessary to connect the given and effective situations – maybe the case for using games in experiments doesn’t need to be stated so strongly anyway?

\*\*\* *Response*: As detailed in our response to the Editor’s first comment, we have again considerably revised the respective paragraph on p. 5. However, if the Editor or Reviewer still deem this paragraph misleading, we are prepared to remove it.

If I understand correctly, the paper asserts i) “payoff” and “utility” are interchangeable terms meaning (quote) “numerical representation of [a players] preferences” p9 ii) “outcomes” means something different, namely the concrete material value that one gets at the end of a game or round, and iii) (quote) “outcomes are understood to directly represent preferences, meaning they directly translate to payoffs”p9 From these three points, is it correctly inferred that outcomes are directly translated to payoffs, and payoffs are identical to utility? If one were to read this from a decision theoretic perspective, could a way to understand this then be to say that in Game Theory, the utility function is always assumed to be linear (with an intercept of 0 and a slope of 1)?

I can see that this might be reasonable in the cases used in the paper, and in the payoff calculations presented. I can also understand that this might be a broad assumption in the Game Theoretic textbooks referred to (except Camerer 2003). So would this be a reasonable framing of the relationship between these concepts, as described in the paper? And is this a reasonable interpretation of what is meant by “preference indicator function”? Some theoretical question remain for me, then. The first concerns the way in which Interdependence Theory is presented as a compliment to game theory. On page 17 it says “whereas in interdependence theory the starting point of any consideration is the given situation, in Game Theory it is the effective situation”. From the way that the given and the effective situations are defined, I read this as saying something roughly like the following (my words): Game Theory analyzes interactive decisions in terms of payoff matrices which represent utility and not outcomes. Interdependence theory compliments Game Theory, because it provides tools to analyze how utilities are psychologically constructed from outcomes, in a way that allows us to capture social motives (and other interesting social psychological processes).

From this understanding, and given the above, would it then be right to understand Interdependence Theory as a kind of expansion or psychological generalization of Game Theory, in which the assumption of a linear utility transformation is relaxed (plus some other conceptual toosl)? Would this then be a reasonable way to read the application of these concepts throughout the paper, and to understand the use of Interdependence Theory to provide psychological meat to the skeleton provided by Game Theory?

This leads to my second theoretical question, which concerns the uniqueness of Interdependence Theory on this point, and whether there aren’t approaches within Game Theory (broadly understood) that aim towards the same objective? The paper explicitly references the strategy, used in Interdependence Theory, of defining social motives in terms of weightings between one’s own and other’s utility functions. This strategy is not unique to Interdependence Theory. Prominent examples include
McClinktock, C. G. (1972). Social motivation: A set of propositions. Behavioral Science, 17(5), 438-455
Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. The Quarterly Journal of Economics, 117(3), 817-869
Janssenn, M. A., & Ahn, T. K. (2006). Learning, signaling, and social preferences in public goods games. Ecology and Society, 11(2).
More broadly, Bayesian and Reinforcement Learning approaches to modeling choice in games also relax the assumption of a linear function relating outcomes and utilities. Examples include
Masel, J. (2007). A Bayesian model of quasi-magical thinking can explain observed cooperation in the public good game. Journal of Economic Behavior and Organization, 64, 216-231; 33;
Camerer, C., & Ho, T. (1999). Experience weighted attraction in normal form games. Econometrica, 67(4), 827-874;
Roth, A., & Erev, I. (1995). Learning in extensive form games: experimental data and simple dynamic models in the intermediate term. Games and Economic Behavior, 8, 164-212.
I think it is this work that has fueled my confusion on this point. Each of the abovementioned papers are cases in which an attempt is made to provide a psychological account of how outcomes are transformed to utilities when playing economic games; or at least the models they present include explicit utility functions for this purpose. And although they may not be canonical examples of classical Game Theory, I would have counted them within the scope of Game Theory (perhaps much) more broadly construed. (Following Camerer 2003, we might call this behavioral Game Theory).

My question then is whether this work also deserves a mention, possibly as a very prominent set of approaches to psychologizing Game Theory? Mention of these approaches could be noted at the beginning of section I (p8), near where it says “In this section we will give an introduction to the two most prominent theories of games…”. At least RL theories might deserve a mention as also being as prominent (both Camerer & Ho and Roth and Erev have >2000 citations). I don’t think a full account of this literature is necessary. That would overcomplicate the narrative. But this at least explains my confusion, and may be a cause of confusion for readers who are familiar with these models, so that some reference to them is helpful.

\*\*\* *Response*: We thank the reviewer for the detailed insights and comments, which helped us further clarify these issues. In the revision, we have implemented the following changes:

* First, the reviewer’s summary of the relation between outcomes, payoffs, and utilities in Game Theory as spelled out in i) and ii) of the reviewer’s comment is right. The quote referred to in iii), however, does not refer to an inherent aspect of Game Theory. Specifically, Game Theory exclusively refers to payoffs and, therefore, does not require any assumptions about how outcomes translate to payoffs. However, the game-theoretic analysis of games as typically used in experiments does assume that outcomes equal payoffs. In the statement the reviewer quoted in iii), we meant to refer to this situation in particular, that is, experiments using game-theoretic paradigms. We now make this explicit on p. 10. Here, we also refer to the linear transformation of outcomes to payoffs as correctly noted by the reviewer as well as to the concept of “preferences” (as included in the term “preference indicator function” the reviewer refers to in his/her comment): “Nonetheless, for the sake of simplicity and in line with prior experimental research (e.g., Bardsley et al., 2010; Bolton, 1991), in all game-theoretic analyses of the games described here, we take the given outcomes as a direct representation of preferences (i.e., higher outcomes are strictly preferred over smaller ones), meaning that the given outcomes directly translate to payoffs or utility, respectively, in a linear fashion.”
* Second, the reviewer is right to assume that this linear transformation of outcomes to payoffs that is often assumed in experimental work using game-theoretic paradigms is relaxed when considering games through the lens of Interdependence Theory. To clarify this, we have added the following sentence on p. 17: “Stated differently, the transformation of outcomes to payoffs (in a game-theoretic sense) may be non-monotonic according to Interdependence Theory.”
* Third, we now make explicit that social motives have already been considered early on by other researchers (even though Interdependence Theory makes the first attempt to integrate social motives in an overarching theoretical framework). In this regard, we also cite the paper by McClintock as suggested by the reviewer (p. 18): “Social motives have early on been considered as drivers of behavior in economic games (e.g., Kuhlman & Marshello, 1975; McClintock, 1972; Messick & McClintock, 1968).”
* Fourth, we have added a footnote on p. 6 in which we refer to other psychological accounts of behavior in games as suggested by the reviewer (and beyond) as follows: “In addition to Interdependence Theory, other theoretical approaches, such as reinforcement learning (e.g., Camerer & Ho, 1999; Roth & Erev, 1995) or generalized reciprocity models (Romano et al., 2017; Yamagishi et al., 1999), have been proposed in an attempt to provide psychological accounts of behavior in games. Here, however, we focus on Interdependence Theory given (i) its broad application to diverse situations of interdependence, (ii) its pioneering role in the field, and (iii) its direct reference to situational features of games and corresponding affordances provided.”

A final point on outcomes and utility is that I do perceive some remaining conceptual slips concerning whether payoff matrices represent outcomes or utilities. This is mainly when introducing the prisoner’s dilemma. For example on page 18 “To illustrate this, again consider the Prisoner’s Dilemma in Figure 1A, assuming that the payoffs in the matrix denote (monetary) outcomes.”

\*\*\* *Response*: We thank the reviewer for pointing us to this ambiguity. To prevent confusion, we have revised the respective sentence on p. 18 as follows: “To illustrate this, again consider the Prisoner’s Dilemma in Figure 1A, now assuming that the values in the matrix denote (monetary) outcomes (rather than payoffs, as would be the case in a game-theoretic consideration).”

It occurs to me that the concept of affordances, as used in this paper, is a component of interdependence theory. Can this also be made clearer when the concept is introduced, and in the conclusion. I think this could pique the readers interest to learn more about this useful application of the concept.

\*\*\* *Response*: Indeed, Interdependence Theory does refer to the concept of affordances. We now make this explicit when introducing the concept in Section I.3: “Several theories in psychology – including Interdependence Theory – consider affordances as key aspects of situations that allow certain dispositional factors to influence behavior” (p. 19).

1. The paper claims in a number of sections that it provides the first theoretical overview of game scenarios. However, there are many book and paper length treatments of this topic. It is also stated that the paper provides the first theoretical introduction to economic games. This is an odd statement, since most of the games described here were developed within applied mathematics, which is purely theoretical, and were not initially developed as experimental paradigms.

\*\* Response: Although we maintain that there is no paper, chapter, or book available that provides a theoretically-founded and practical introduction into games that is specifically directed at experimental psychologists, we no longer imply by our wording that we are the first to provide such a summary.

\*\*\*Reviewer: I would argue that this is one of the goals of Camerer 2003, especially the earlier chapters, which give a detailed and accessible introduction to the dictator and ultimatum games (the content of this presentation is not unlike the one given in the present manuscript). Camerer also works to provide the theoretical terminology (and formal tools) for other social scientists. This is arguably also the goal of e.g. Coleman 1995 – “Game Theory and It’s Applications in the Social and Behavioral Sciences”. Could the paper perhaps be situated as a contribution to this kind of introductory work, that updates (or focuses on) Interdependence Theory?

\*\*\* *Response*: We thank the reviewer for hinting us at these textbooks. To acknowledge them in our paper, we have added the following footnote on p. 4: “We do want to emphasize that there are entire textbooks available that introduce economic games for experimental scientists (e.g., Camerer, 2003; Chaudhuri, 2009; Colman, 1995). Here, we specifically aim at providing a more hands-on and theory-driven overview that also refers to Interdependence Theory as well as the affordance concept.”

1. There is a conceptual confusion about how to interpret behavior within classic game theory. In the introduction the authors present game theory in terms of mental concepts and events. For example, they say: “Game Theory provides a formal analysis of games and behavior in games that takes into account all players’ preferences (i.e., their attitudes towards the distribution of payoffs between themselves and others)”. However, in game theory, games are mathematical objects, and equilibria etc are defined using mathematical proofs, so the allusion to psychological states here seems inappropriate.

\*\* Response: We apologize for the ambiguous wording. To address this issue, we have considerably revised the respective paragraph on p. 7, now strictly sticking with game-theoretical terms without alluding to attitudes or other psychological concepts.

\*\*\*Reviewer: I still see some tension here. For example, it is my understanding that the “effective situation” refers to something like a psychological interpretation of the game, e.g. (quote) “players are assumed to engage in a psychological transformation, which eventually results in an effective situation” on p16. In this context the paper also says (quote) “whereas in Interdependence Theory the starting point of any consideration is the given situation, in Game Theory it is the effective situation” p17. This is a more minor issue, so I don’t think anything needs to be changed, but I think it is also connected to the broader issue raised above concerning utility etc.

\*\*\* *Response*: Given the changes we made in response to the reviewer’s (second) comment above, we hope that the tension the reviewer refers to here is now also solved.

1. The title could be improved. There are no formal models provided in the paper, and the focus is much broader than prosociality.

\*\* Response: As suggested, we have changed our title to “Economic games: Modeling social interactions and understanding prosocial behavior”. We hope that, this way, the title better conveys that games are sought to specifically model different kinds of social situations, but not to formally model prosocial behavior.

\*\*\*Reviewer: There may be readers for whom “model” strongly cues formal theory, but I leave this entirely to the authors’ discretion.

\*\*\* *Response*: We thank the reviewer for making us aware of potential confusion among readers concerning the term “model”. We have thus again changed the title to “Economic games: An introduction and guide for research”.

**Final Editor decision—Accept**

Jan 28, 2021

Dear Dr. Isabel Thielmann,

I have now had a chance to read over your manuscript “Economic games: An introduction and guide for research”, along with the letter describing the changes you made. Thank you for your responsiveness to the concerns that the reviewers and I raised. I am happy to say that your paper is now officially accepted for publication in Collabra: Psychology. Congratulations on this excellent work, I think it 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.

I’ll just take this opportunity to point out that I think that your response to my point 1 clears up the issue mostly. I will add here that if, as your paper states, “preference is the construct of interest one aims to explain” then your point about the benefits of economic games not making claims about unobservable constructs stands and I agree. And this is indeed a benefit of economic games. But your framework goes beyond this. You state in your response that the goal of the paper is “to show how an affordance framework can help illuminate the (otherwise unobservable) psychological processes involved in prosocial behavior”. This is now about unobservable psychological processes. In fact, the framework you propose is about drawing inferences about people’s unobservable psychological states. So, in effect, this benefit of the games described above is irrelevant as far as your framework is concerned. Having said that, I don’t think the strength of your paper is impacted much by this paragraph. Nonetheless, I think that the new wording of the paragraph make things clearer to me and probably to other readers.

As there are no further reviewer revisions to make, 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, Farid Anvari