Step 1 Evaluation Report

CONFIDENTIAL

Call reference	ERC-2021-STG
Activity	HORIZON ERC Grants
Funding scheme	ERC STARTING GRANTS
Panel name	SH4
Proposal No.	101042485
Acronym	SUSHI
Applicant Name	Eric SCHULZ
Title	Scaling Up Studies of Human Intelligence

EVALUATION CRITERIA

Criterion 1 - RESEARCH PROJECT
Ground-breaking nature and potential impact of the research project
To what extent does the proposed research address important challenges?
To what extent are the objectives ambitious and beyond the state of the art (e.g. novel concepts and approaches or development between or across disciplines)?
To what extent is the proposed research high risk/high gain (i.e. if successful the payoffs will be very significant, but there is a high risk that the research project does not entirely fulfil its aims)?

Scientific Approach
To what extent is the outlined scientific approach feasible bearing in mind the extent that the proposed research is high risk/high gain (based on the Extended Synopsis)?

Criterion 2 - PRINCIPAL INVESTIGATOR
Intellectual capacity and creativity
The questions below can have one of the following five responses: Exceptional/Excellent/Very Good/Good/Non-competitive
To what extent has the PI demonstrated the ability to conduct ground-breaking research?
To what extent does the PI provide evidence of creative independent thinking?
To what extent does the PI have the required scientific expertise and capacity to successfully execute the project?

PANEL SCORE AND RANKING RANGE

Final panel score: B (is of high quality but not sufficient to pass to Step 2 of the evaluation. Please note that you may also be subject to resubmission limitations in the next call)	**Ranking range*:** 24%-33% For your information, only the top 23% of the proposals evaluated in panel SH4 were retained for Step 2.

* Ranking range of your proposal out of the proposals evaluated by the panel in Step 1, in percent, from 1% for the highest ranked proposals to 100% for the lowest ranked.

PANEL COMMENT

This evaluation report contains the final score awarded by the ERC review panel during the first step of the ERC Starting Grant review and the ranking range. The discussion of the panel was conducted within the context of the individual reviews submitted by ERC panel members.

The panel closely examined all the individual review reports and, while not necessarily subscribing to each and every opinion expressed, found that they provide a fair overall assessment. The comments of the individual reviewers were the basis for the discussion and the final recommendation of the panel, and are included in this report.

This is an exceptional PI: the PI has published key papers, has won several awards, and is now leading their own Research group.

The panel also agreed that the proposal is ambitious and potentially high risk/high gain. However, its details are not worked out well. Most importantly, the theoretical and conceptual framework needs extra work. Related to this, the PI should further justify whether the proposed games (and the overall approach more generally) are indeed appropriate to study human intelligence. Therefore the panel reached the conclusion that the subject was timely and interesting, but the proposal was not mature enough and further work on it is necessary.

Overall the panel considers this proposal to be of good quality. However, based on the combined set of criteria used in the assessment it was not ranked highly enough to be retained for Step 2. The panel therefore recommends that the proposal should not be retained for Step 2 and should not be considered for funding.

REVIEWER COMMENTS

The following individual reviews have been carried out independently prior to the panel meeting and do not necessarily reflect the panel's final opinion

Reviewer 1

Research Project
Ground-breaking nature and potential impact of the research project The proposal is based on the excellent idea to investigate exploration, as a core function of intelligence, with rich and large-sample data from online games. It presents a comprehensive and innovative approach to several important aspects of exploration and provides strong theoretical background for studying those. What is missing in the theoretical argumentation, though, is the extensive extant psychometric work on intelligence, which has demonstrated the very high relevance of reasoning (fluid intelligence) for all kinds of (academic, occupational, everyday problem solving) outcomes. The proposal therefore rather puts forward interesting aspects regarding exploration behavior than truly “scaling up studies of human intelligence” – which would require linking the present ideas to existing models of working memory and reasoning and how they interact with declarative and procedural knowledge in skill acquisition and complex problem solving. **Scientific Approach** The specific approaches used to implement the different work packages are somewhat underspecified in B1. The choice of games and the way they are adapted seems very appropriate, but one has to trust in the PI’s ability to develop the necessary complex models based on the rough sketches in B1. This is particularly the case for Aim 3, where the collaborative exploration adds a challenging level of complexity.

Principal Investigator	
To what extent has the PI demonstrated the ability to conduct ground-breaking research?	Exceptional
To what extent does the PI provide evidence of creative independent thinking?	Exceptional
To what extent does the PI have the required scientific expertise and capacity to successfully execute the project?	Exceptional

Comments (Optional for reviewers)
The PI has an outstanding track record on all relevant dimensions, with extensive funding, excellent publications, a great network of collaborators, and comprehensive education in psychology, statistics, and computer science.

Reviewer 2

Research Project
Ground-breaking nature and potential impact of the research project SUSHI is a project at the interface between psychology and machine learning; it raises interesting problems, and promises to lead to novel and ground-breaking outputs. The PI starts from observing that while researchers in machine learning want “to build artificial systems that exhibit human-like abilities, but do not study the underlying principles of intelligence driving these abilities”, psychologists “could provide insights into the principles of human intelligence, but focus on superficial paradigms and apply simplified models.” (p. 1) Obviously, the PI wants both sides to learn from one another, but this is puzzling given his own description of the psychology of intelligence: if psychological research has produced only “superficial paradigms and … simplified models”, then what good could they deliver to research in machine learning? Or does the PI say that, in principle, psychologists should deliver the principles, only as a matter of fact they just haven’t? If the latter, then how “could they provide insights into the principles of human intelligence” after all? However this may be, the PI wants to use computer games in order to study a “key feature of human intelligence: the ability to learn and explore efficiently” and proposes “to study 3 principles of exploration: compositionality, empowerment, and collaboration.” (p. 1) These three principles are used to structure the whole research project. The PI is highly optimistic that computer games are a “a realistic testbed to investigate human exploration behavior”, and that one can ignore that real-life search for things often acts under constraints such as that one has no clear rules of the game, has to take care of other tasks as well, but can also think outside of the box (which I assume one cannot do in most computer games). If I explore a sport, such as tennis, is how I try out various shots and their effects on reaching my goals really governed by the same principles as when I learn to play computer tennis? The PI is right that in both “real” and computer activities, say, compositionality plays a huge role, but I suspect that it does so in very different ways. Someone who can be a great real tennis player can fail miserably at computer tennis, and vice versa. Wouldn’t that be important to understand first? Other problems are conceptual. The PI states “…efficient exploration requires multiple ingredients of intelligence including learning, abstraction, and planning: exploration is at the core of intelligent behavior.” (p. 3) What does “core” mean here? A necessary or a sufficient condition, or just a strong association? Some standard intelligence tests do not involve tasks requiring exploration at all; they just ask whether one knows a certain rule, say, of geometry, language, or logic and how fast one is in applying the rule correctly. In addition to this problem, I do not see a serious discussion of the concept of intelligence in the project. Such problems with some starting points make it difficult to see a clear high-risk-high-gain aspect in the work. For instance, while I do not doubt that composition, empowerment and collaboration are helpful components of successful goal-directed behavior, I do not see that as leading to a theory of intelligence; nor does the PI state what exact principles he thinks might govern these components. The PI describes as risky, for instance, the handling of large data (p. 6), but that seems more like a managerial difficulty, not an epistemic risk. **Scientific Approach** The scientific expertise for such a project is in principle there: the PI is fully immersed in psychology and computer science. The project looks feasible, given the PI’s preparation and publications on several of the discussed topics. But, as stated, I am underwhelmed by its conceptual and theoretical sophistication, and I do not see a bold or risky hypothesis.

Principal Investigator	
To what extent has the PI demonstrated the ability to conduct ground-breaking research?	Good
To what extent does the PI provide evidence of creative independent thinking?	Very good

To what extent does the PI have the required scientific expertise and capacity to successfully execute the project?	Excellent

Comments (Optional for reviewers)
The PI is obviously successful in his discipline(s), having won several awards and grants, directing several PhD students, and other things beyond. There is little doubt that he can direct a project on machine learning and computer games. I am not convinced that he can pursue a theoretically ambitious interdisciplinary project.

Reviewer 3

Research Project
Ground-breaking nature and potential impact of the research project The project proposes to “upgrade” studies concerning human intelligence in such a way that specialists in machine learning can use these more sophisticated studies to compare human and machine intelligence. The upgrade that is being proposed is the study of human behaviour while involved in computer games. It is safe to assume that the worlds of computer games are indeed more complex and challenging than the standard tasks put to subjects in laboratory conditions. So it is a sort of intermediate level that could forge links between human and machine intelligence. This research is broken down in three components: compositionality, empowerment, and collaboration. This leads to three subprojects. These are presented in the application to a certain degree of detail, which allows to deduce that the tasks are indeed challenging but they seem feasible. As indicated in the application itself, the risks are very high indeed. That a high gain is involved, seems clear. If it is indeed the case that there is such a gap between human and machine intelligence then a lot is to be learned from any study to aims at bridging this gap. If I look at the size of the case studies proposed, the PI has set the standards very high indeed. There are two things that seem worrisome: - In the more detailed description of the tasks, use is made of computer software that itself involves forms of learning. Is there possibly a form of circularity at work here and is this problematic? - The PI mentions that one should assume that the three aspects - compositionality, empowerment, and collaboration – are not independent from one another and therefore the question must be raised: how will this be controlled? The application provides some comment but that does not really clarify the matter. **Scientific Approach** It is not easy to judge the feasibility of this project for the following reasons: - There is as good as no information about the team members and the distribution of the tasks over the team members. This is important for it is an entirely different situation if team members are assigned to one of the three subprojects or instead to all of them being involved in all the tasks, - It is mentioned that one of the great challenges is to handle the amount of data that will come out of the three tasks. Does that mean that it is unclear at this moment whether that will be possible or not? That is truly a great risk to take and to have one specialist in big data on board does not really sound reassuring, - Perhaps this need not be mentioned as such, but what infrastructure is required to execute these tasks?

Principal Investigator	
To what extent has the PI demonstrated the ability to conduct ground-breaking research?	Exceptional
To what extent does the PI provide evidence of creative independent thinking?	Exceptional
To what extent does the PI have the required scientific expertise and capacity to successfully execute the project?	Excellent

Comments (Optional for reviewers)
- The “standard” elements need not be repeated: attending conferences, organizing workshops, international collaborations, teaching, supervision of master students and PhD students, etc. These are all present in acceptable amounts for a researcher that obtained a PhD in 2017, - Of special mention is the fact that the PI is head of a research group at the Max Planck Institute, - There are four grants obtained, either as PI or as co-PI, and no grants submitted. None of these grants are related to this project, - The key publications are all in highly-ranked journals, basically the top journals in the PI’s domain.

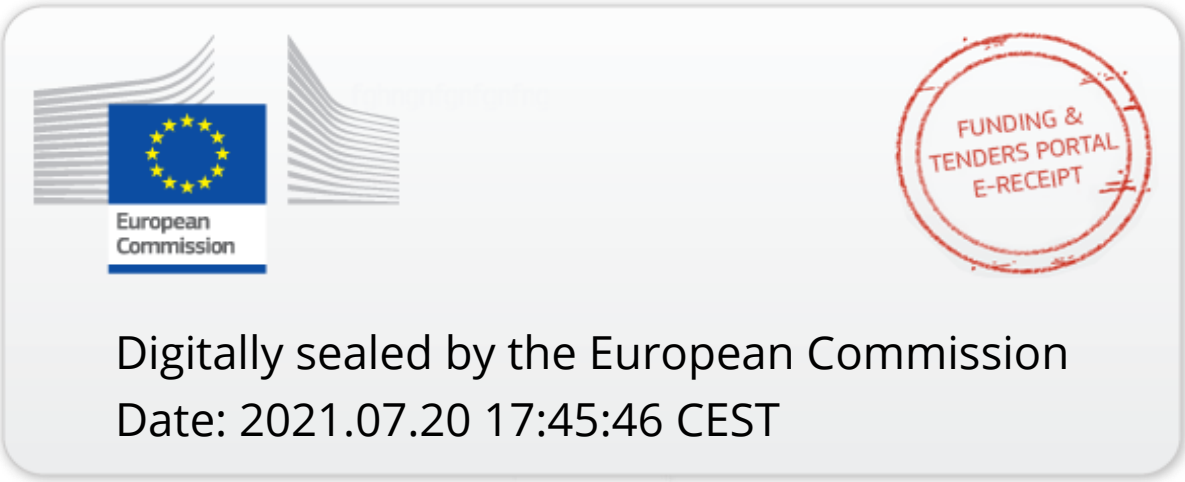
It seems clear that if anyone can do the research proposed in this application, the PI is the right candidate for it.

Reviewer 4

Research Project
Ground-breaking nature and potential impact of the research project This project is certainly ambitious, beyond the state of the art, and high risk/high gain. The only quibbles I have is whether the skater (complex motor skill) game used for Aim 1 really taps into in intelligence? How would this be modelled (Aim 1)? And related to the first point: is it really appropriate to place exploration in such a specific context at the core of intelligence (although this might be somewhat of a semantic issue)? **Scientific Approach** The project will undoubtedly come with many challenges/risks, but the gains could be high. The candidate seems to be ideally placed to (at least try to) tackle these challenges. It was not entirely clear to what extent the data were already collected (they still have to be collected for the first two aims but not for the third one?), and if they still have to be collected, how quickly this could be done.

Principal Investigator	
To what extent has the PI demonstrated the ability to conduct ground-breaking research?	Exceptional
To what extent does the PI provide evidence of creative independent thinking?	Exceptional
To what extent does the PI have the required scientific expertise and capacity to successfully execute the project?	Exceptional

Comments (Optional for reviewers)
No comments received



This electronic receipt is a digitally signed version of the document submitted by your organisation. Both the content of the document and a set of metadata have been digitally sealed.

This digital signature mechanism, using a public-private key pair mechanism, uniquely binds this eReceipt to the modules of the Funding & Tenders Portal of the European Commission, to the transaction for which it was generated and ensures its full integrity. Therefore a complete digitally signed trail of the transaction is available both for your organisation and for the issuer of the eReceipt.

Any attempt to modify the content will lead to a break of the integrity of the electronic signature, which can be verified at any time by clicking on the eReceipt validation symbol.

More info about eReceipts can be found in the FAQ page of the Funding & Tenders Portal.

(https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/support/faq)