Choosing Theories Matters: How Different Learning Theories Impact Simulation Results Using Equine Stereotypies as Example

Veronika Kurchyna[1,2,3], Jan Ole Berndt[2], and Ingo Timm[1,2]

[1] Trier University, Business Informatic and AI I
[2] German Research Center for Artificial Intelligence, Cognitive Social Simulation
[3] Corresponding Author: kurchynav@uni-trier.de

Abstract. Cognitive Social Simulation combines agent-based models and psychological theories to explore the mechanisms behind individual behaviour and emergent social phenomena. We examine the development of obsessive-compulsive behaviours in horses to show how the choice of learning mechanism impacts derived policy recommendations. The experiments compare self-reinforcing operant conditioning and social learning both at baseline and with added interventions for welfare. Our results showcase how experiments with two mechanisms and similar baseline calibration are not comparable and stress the importance of expert validation in choosing psychological mechanisms during modelling.

Keywords: Cognitive Social Simulation · Reinforcement Learning · Equine Stereotypies · Learning Mechanism Comparison · Animal Behaviour

1 Introduction

Horses have been domesticated companions to humans for thousands of years. Due to their complex social structures and emotional intelligence, some research has even gone as far as to drawing parallels between equine and human psychology[11]. Besides animal comparisons, humans also resort to computational methods to explore and explain observations. Cognitive Social Simulation combines agent-based simulations with psychological theories to explore the mechanisms behind individual behaviour and emergent social phenomena [10]. Often, different theories compete to explain the same reality, without conclusive evidence on the validity of one over the other [26]. However, the chosen theory (and even its formalisation) [17] has major impact on the outcome of a simulation study, which is critical when results serve as decision-support for the development of policies. Equine behavioural medicine (the study of behavioural issues such as compulsions or aggression) offers an unusual use case in which empirical evidence discredits one theory in favour of another: Research found no evidence of social learning in horses regarding tasks or undesirable behaviours, but confirms operant conditioning as the primary learning mechanism for equids [8].

To show the profound impact of theory choice on the behaviour of agents, we propose a model of the development and spread of equine stereotypies, undesirable coping mechanisms for repeated stess or prolonged boredom [19]. It is a common misconception among horse owners that stereotypies are acquired through observation, despite empirical evidence suggesting otherwise [1]. This contradiction demonstrates how different learning modes can lead to comparable observations at baseline and diverge once additional components, such as behavioural interventions, are introduced. Building on previous work (e.g.[16]), we show how interventions can have wildly different outcomes depending on the chosen learning mechanism despite similar baseline model behaviour.

2 Background

Agent-Based Social Simulation (ABSS) with autonomous artificial agents representing living entities with complex social behaviours and cognitive processes, typically human actors. In contrast to other methods, ABSS is a theory-driven approach with emphasis on transparency, interdisciplinarity and underlying mechanisms even in the absence of empirical data specific to the use case.

Behaviour (such as compliance) plays a significant role in the outcome of many system-level interventions. With competing psychological theories explaining the same reality, modellers face uncertainty on the selection, interpretation and formalisation of concepts [2]. As an interdisciplinary field, social simulation requires both domain knowledge and expertise on the modelled theories.

2.1 Learning by Doing and Learning by Seeing

While the exact definition of learning is controversial, it is generally agreed-upon that learning encompasses the cognitive and subconscious processes leading to behavioural modification based on environmental cues [23].

Similar to [16], this work focuses on two mechanisms: *Self-Reinforcing Operant Conditioning* and *Social Learning*. In Self-Reinforcing Operant Conditioning, actions are reinforced through internal rewards not administered by another party, as observed in different animals [9]. In horses, spontaneous (inter-)actions with(in) the environment lead to learning opportunities, such as opening stall doors and reach new food sources by randomly mouthing at the latch [15].

In contrast, *Social Learning* is not self-contained: reinforcement does not only happen upon execution, but also through observation, provided a mental link between action and outcome is established by the observer [3]. The acquisition of complex behaviours and rituals is documented in different mammalian and avian species as a a potential evolutionary advantage and driver of social structure [12].

On a technical level, the *q-Learning*-Algorithm serves as formalization of operant conditioning [24], with the utility of actions being state-dependent. To include social learning, the *q-values* of a State-Action pair can be influenced by the observation of this action. A well-designed definition can fulfill the social criterion by the inclusion of relevant characteristics that establish sufficient similarity for learning to take place [20].

2.2 Equine (Mis-)Behaviour

Like most domestic animals, horses depend on humans to care for their wellbeing [7]. Besides basic physiological needs, horses are highly social animals, showing a variety of social behaviours such as those shown in Fig. 1.

In case of animal abuse or prolonged periods of unmet needs, horses develop coping strategies. Besides obvious misbehaviour, such as aggression towards humans or other animals, subtle coping strategies can cause long-term damage [4]. These mechanisms are called *stereotypies*, defined by habitual execution - even without obvious trigger and often persistent after the stressors have been removed. Two maladaptive coping-mechanisms are of particular importance due to their frequency and potential for long-term harm: **Cribbing** (or windsucking) horses bite down on a hard surface (such as wood or metal) and swallow air, often in response to malnourishment[19]. Besides damage to both teeth and environment[4], the ingestion of air greatly increases the risk of lethal colics and other digestive issues. **Weaving** is less destructive - horses sway from side to side on their front legs, causing serious wear and tear on their hooves, joints and muscles, and increasing the long-term risk of developing conditions affecting mobility. This behaviour has an association with isolation[18].

Stereotypies make for an interesting use case due to the these characteristics:

1. The underlying physiological and emotional processes and the effects of treatments and preventive measures for stereotypies are well studied.
2. Previous work successfully established learning as explanatory model for the development of behavioural anomalies [25].
3. Once acquired as habitual coping mechanism, horses are unlikely to stop and will continue to display these behaviours even without prior trigger [13].
4. Empirical research provides no evidence for social learning - both in training contexts and stereotypy-acquisition in stabled horses[22].
5. Despite the scientific evidence, numerous horseowners believe that stable vices are acquired through observation and imitation [1].

Fig. 1: Visual examples of horses engaging in (a) allogrooming, (b) cribbing and (c) normal interaction across fences. Pictures taken by author.

Fig. 2: Overview of the environment with individual stalls, two riding rings, three pastures and human-restricted pathways.

3 Model Description

The **grid-environment** is a barn in which horses are kept stalled with regular turnout (pasture time) and human interaction (grooming, riding and changing locations). **Horse-Agents** are characterised by attributes including stress level, loneliness, boredom and learned behaviours as state-action-utility table as used in reinforcement learning. The grid cells differ in category (stall, pasture, arena, human passageway) based on the sketched map in fig. 2. Each cell ($9m^2$) can be occupied by one horse at a time. Horses' ability to move and interact is limited by location: they cannot leave stalls or connected areas (such as pastures or riding rings) on their own. **Humans** are *exogenous* factors that can move horses to another cell, add or remove enrichments or stressors, administer punishments or restrictions for unwanted behaviours, or handle them during different activities.

3.1 Scheduling and Activities

Each simulation step represents 30 minutes. Based on equine schedules[5], horses always perform one of four major activities (see Fig. 3), with possible subactions during the overarching activity. The frequency and duration of different handling actions, turnout and initial distribution of stereotypies are based on empirical data. While the major action is a set schedule for all horses, these subactions during activities are the objects of learning: Horses can **interact** with others while on a pasture or in a stall with occupied adjacent stall durin standing hours. **Observation** of others within range (c.f. 3.2) is possible at all times except during lying phases. **Stereotypies** can only be performed while eating or standing - not while lying down or being handled by a human.

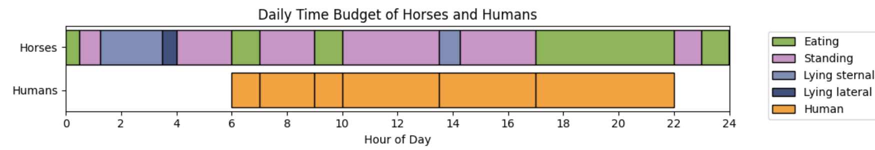

Fig. 3: 24 hours of equine and human activity time windows based on [5]

3.2 Social Interactions and Perceptions

Horse-Agents have three different ranges for perceptions [6]. Within these ranges, their moods are affected by the presence of other horses, with lack of social contacts greatly increasing the stress-level.

1. *Awareness*: Within approximately 200m, a horse is aware of the presence of another in its vicinity with no perception of activities or state.
2. *View:* If unobstructed by walls, horses have a view distance of 3 cells (9m). Within this range, other horses can be observed regarding their activities, state and moods, which is required for observational learning.
3. *Interaction*: Within 3m, direct interaction is possible, both allogrooming and aggression. This range is expressed using the Moore Neighborhood [21].

3.3 Learning

To model learning as described in 2.1, simplified q-learning was implemented to map the current state S(tuple of activity, place, stress and other variables) and available subactions (i) to continuously updated utilities $(u(S,i))$. The binary probability p_i for a horse to execute i depends on $u(S,i)$ and the constant utility threshold m_{ut}, which is arequired for execution, similar to the multi-action reinforcement learning approach presented by [16]. Execution can also happen if i is a habitual action ($h_i = True$). Additionally, random execution can occur if a random number r (between 0 and 1, sampled using normal distribution) is below the threshold defined by the exploration-rate m_{ex}, modified by stress-level s - higher stress levels increase the probability of random attempts at relieving this stress. Thus, the decision of execution is made as follows:

$$p_i(S) = \begin{Bmatrix} 1 \; \textit{if} & u(S,i) \geq m_{ut} \\ \quad \textit{or} & h_i = \text{true} \\ \quad \textit{or} & r \leq m_{ex} \times \frac{1}{s} \\ 0 & \text{otherwise} \end{Bmatrix} \tag{1}$$

The reinforcement of the action is the stress reduction sr_i adjusted by the learning-rate m_{lr} and s, leading to diminishing returns of cribbing and weaving for relaxed horses. As such, the reward for actions is endogeneous, turning this into *self-reinforced* conditioning. Additionally, frequently stressed horses eventually start executing i habitually - once a known state's utility of an action reaches above a multiple of m_{ex} (in our calibration, 1.5 was found as threshold value), the corresponding h_i is set to *True*.

In that case, the behaviour is frequent enough to be observable for other horses. This is required for social contagion, which happens for each observable Horse e in observation range H_{OR}. The reinforcement for action i depends on the social learning rate m_{slr}, since the true effect of an action cannot be observed directly. For $m_{slr} = 0$, no observational learning takes place. Thus, the knowledge matrix entry corresponding to the observed horse's state is updated as follows:

$$\forall e \in H_{OR}, \text{ if } h_i(e) \implies u(S_e, i)_{new} = u(S_e, i)_{old} + m_{slr} \tag{2}$$

3.4 Interventions

Equine Stereotypies are difficult to treat. While some methods, such as medication (SSRIs in particular) are possible, they are excluded from this model since scientific evidence is largely still work in progress. Therefore, we only consider non-pharmaceutical interventions, whose effectiveness is linked with added stress reducing the reward of the behaviour, rather than the development of positive coping mechanisms. There is also the possibility of adding enrichment, although horses primarily only interact with novel items upon initial exposure [14].

Each intervention c is associated with a probability p_c of being applied to the horse or its stall. The reward of performing i is decreased by the parameter $m_{p(c)}$ as positive punishment.

Since research indicates that such interventions present no long-term improvement of unwanted compulsions, the primary means of reducing and preventing stereotypies are improving living conditions, namely turnout and social interaction. While the model is set up to support different types of interventions in future expansions, our experiments focus on the most promising option: improved welfare through increased time on pastures with other horses.

The increase of turnout time is an important instrument not only for equine welfare, but also for the examination of the research question: if stereotypies are results of social contagion, the improvement of living conditions will not lead to a decline in newly acquired stereotypies, since welfare alone is not the only determinant in the case of social learning. Therefore, the effectiveness of this intervention depends on whether horses learn through observation or not.

4 Experiments and Evaluation

We hypothesize that in the presence of social learning, individual measures for increased animal welfare will show less impact on the development of behavioural anomalies than a setup without social learning. This assumes that if the erroneous belief of social contagion of stereotypies is true, then improved welfare would lose effectiveness due to the influence of already affected horses.

This model is an abstraction of reality, where experimentally determined values yielded from calibration are sufficient for demonstrative purposes. For actual studies, such as policy recommendations regarding animal welfare, the authors strongly recommend consulting the available detailed data on the neurochemical processes and biological markers of different behaviours for precise simulation[4].

Two baseline parameter configurations were identified: the *default* configuration (P_d) sets m_{slr} to 0, while *alternative* configuration (P_a)includes m_{slr} as calibration parameter. Both configurations aim to achieve an approximate distribution of 5% habitual cribbers and 8% habitual weavers. As such, both configurations lead to similar observations.

Figure 4 displays the average stress levels of horses as well as the percentage of cribbers/weavers over the course of a simulation run. While, in reality, such habits form over the course of years, the model was adjusted to yield the same observations within a significantly shorter timespan.

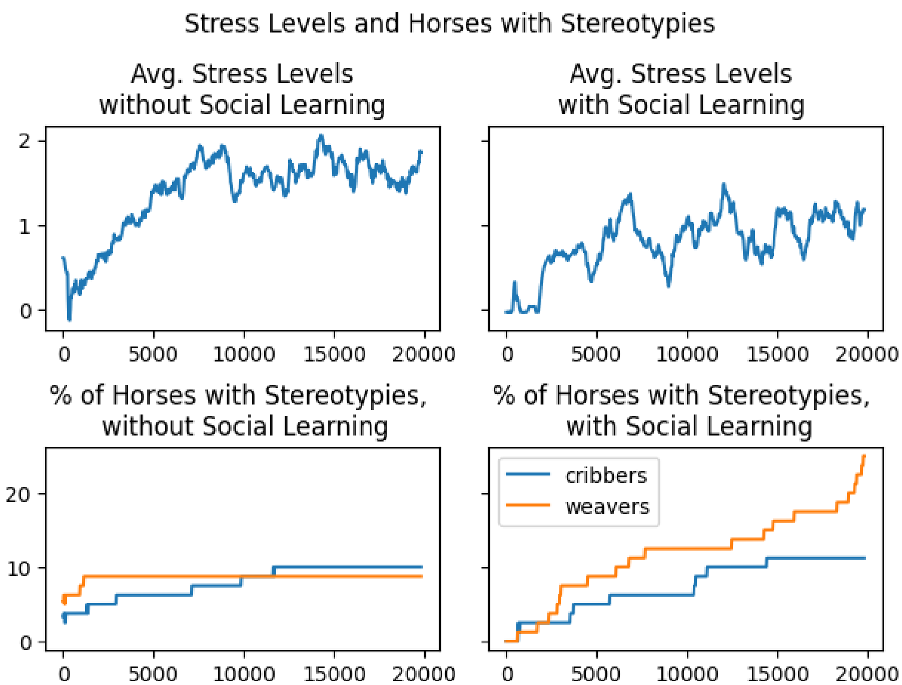

Fig. 4: Stress Levels and Percentage of Stereotypies (average of 30 iterations)

The hypothesis guiding our experiments is *If social learning is available, increasing the chance of longer turnouts impacts the rate of stereotypies at the end of the simulation less than without social learning.* For that, we adjust the model parameter $m_{p(pi)}$ to increase the likelihood of a human permanently prolonging pasture time for a horse by 30 minutes (one simulation step).

To the authors' surprise, this hypothesis is not actually confirmed by the model results, which indicate a more complicated relationship between learning modes and intervention strength. When comparing the effectiveness of the two different parameter configurations, one can see a steady decline of stereotypy cases for P_d, while the results for P_a display a non-linear relationship between behavioural issues and $m_{p(pi)}$.

In fact, the nonlinear relationship is even more apparent for larger m_{slr}, as shown in 5. While the overall pattern remains the same, the magnitude of differences is much greater than anticipated. The expected outcome would be similar levels of stereotypies across the different levels of $m_{p(pi)}$.

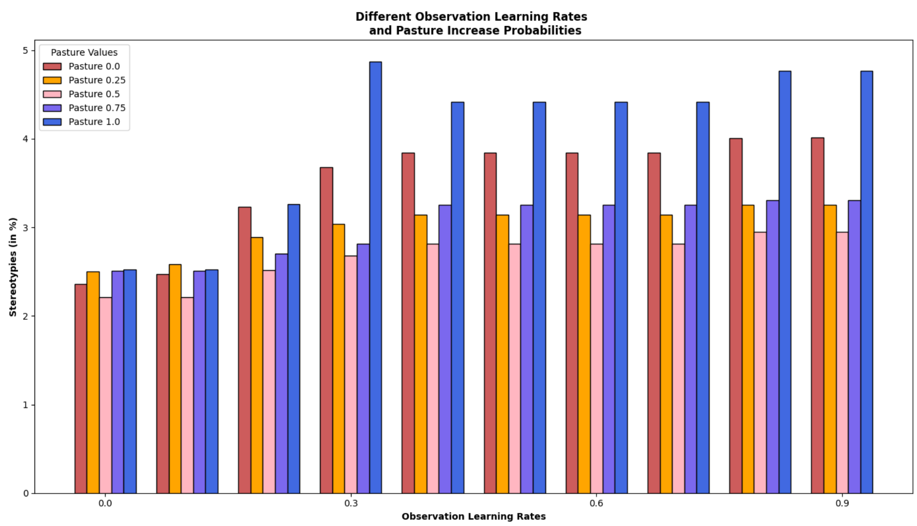

Fig. 5: Rate of Stereotypies based on m_{slr} and $m_{p(pi)}$.

These results, while unintuitive, are not unreasonable. With more social interactions on pastures, increased pasture time does not only provide stress relief from confinement, but also more opportunities for observational learning. In this regard, increasing turnout would be actually *harmful* in regards to obsessive-compulsive behaviours. While plausible within the logic of the model variant with social learning included, this behaviour is not supported by empirical studies. In the absence of contradictory evidence, the plausibility of this observation can lead to blind trust in the findings of the model.

To control against the possibility of the overall parametrization P_a causing this observation pattern, the same experiment was performed on P_d with the inclusion of different m_{slr} in the parameter space. With an occurrence of stereotypies spanning from 10 to 50%, the results are not comparable to the $2-5\%$ observed for P_a. While the baseline behaviour, as shown in figure 4, for P_d and P_a is comparable with similar patterns of stress and stereotypies in terms of shape and magnitude, experiments with interventions no longer yield meaningful results when exchanging the two configurations.

5 Conclusion and Further Work

Social Simulation is applied across different domains, often meant to guide experts and decision-makers by providing informed assessments of the impact of different actions. The field has long moved beyond asking *if* psychological theories should guide modelling efforts and now asks *which* theory to choose, with many competing approaches explaining the same phenomenon.

While horses who display obsessive-compulsive behaviours is an unusual use case for social simulation, this is a real-world example that demonstrates how different theories might explain the same observations but lead to different results once interventions are introduced. Moreover, the changes might even be difficult to predict, as showcased by the unexpected results in Fig. 5 which defied the initial hypothesis. In fact, the conclusion to be drawn using P_a would even be detrimental to animal welfare, if this model were to be used to derive policy recommendations. While the explanation of social learning as driving factor appears logical, the model assuming such a learning mode yields plausible results counter to empirical evidence regarding equine behaviour and animal welfare.

In this use case, reality is well explored, with empirically proven guidelines and studies that give us insight into the cognitive mechanisms at work. However, for humans such data often does not exist, with different learning mechanisms and cognitive architectures competing to explain the same observations without conclusive evidence in terms of validity. Therefore, this work showcases the need to investigate the impact of modelling choices not only on baseline behaviour, but also on the effect of interventions. For further work, the authors intend to examine a use case with human subjects to investigate whether the same phenomenon can be reproduced in an agent-based model of a human population.

ACKNOWLEDGEMENTS

The authors thank the horses Lord Parlando, Toothfairy Luke and Luna for the inspiration for this research. We also thank our human colleagues Nhat Tran, Christian Lohr and Ye Eun Bae for their support during the model preparation and the finalisation of the draft.

References

[1] J. D. Albright et al. "Crib-biting in US horses: Breed predispositions and owner perceptions of aetiology". In: *Equine Veterinary Journal* 41.5 (2009), pp. 455–458.

[2] A. Bandura. *Psychological modeling : conflicting theories / ed. by Albert Bandura ...* eng. Chicago, Ill: Aldine-Atherton, 1971.

[3] A. Bandura. “Social Learning Theory”. In: *Encyclopedia of criminological theory* 2 (2010), pp. 66–72.

[4] B. V. Beaver. “Equine Behavior Problems”. In: *Equine Behavioral Medicine*. Ed. by B. V. Beaver. Academic Press, 2019, pp. 311–345.

[5] B. V. Beaver. “Equine Locomotive Behavior”. In: *Equine Behavioral Medicine*. Ed. by B. V. Beaver. Academic Press, 2019-01-01, pp. 237–277.

[6] B. V. Beaver. “Equine Social Behavior”. In: *Equine Behavioral Medicine*. Ed. by B. V. Beaver. Academic Press, 2019-01-01, pp. 115–150.

[7] B. V. Beaver. “Equine Welfare”. In: *Equine Behavioral Medicine*. Ed. by B. V. Beaver. Academic Press, 2019, pp. 347–369.

[8] B. V. Beaver. “Learning”. In: *Equine Behavioral Medicine*. Ed. by B. V. Beaver. Academic Press, 2019, pp. 65–95.

[9] B. Brembs. “Spontaneous decisions and operant conditioning in fruit flies”. In: *Behavioural Processes* 87.1 (2011). Society for the Quantitative Analyses of Behavior, pp. 157–164.

[10] N. Gilbert. “Emergence in social simulation”. In: *Artificial Societies* (1995).

[11] M. Hausberger et al. “Could Work Be a Source of Behavioural Disorders? A Study in Horses”. In: *PLOS ONE* 4.10 (2009-10-28). Publisher: Public Library of Science, e7625.

[12] C. Heyes and B. Galef. *Social Learning in Animals: The Roots of Culture*. 1996.

[13] B. Hothersall and R. Casey. “Undesired behaviour in horses: A review of their development, prevention, management and association with welfare”. In: *Equine Veterinary Education* 24.9 (2012), pp. 479–485.

[14] G. H. M. Jørgensen, S. H.-O. Liestøl, and K. E. Bøe. “Effects of enrichment items on activity and social interactions in domestic horses (*Equus caballus*)”. In: *Applied Animal Behaviour Science* 129.2 (2011-01-31), pp. 100–110.

[15] E. L. Krueger K and B. R. “Animal behaviour in a human world: A crowdsourcing study on horses that open door and gate mechanisms.” In: *PLoS ONE* 14.6 (2019).

[16] V. Kurchyna et al. “Learned Futility: How Social Learning Can Lead to the Diffusion of Ineffective Strategies”. In: *Social Computing and Social Media: 16th International Conference, SCSM 2024, Held as Part of the 26th HCI International Conference, HCII 2024, Washington, DC, USA, June 29–July 4, 2024, Proceedings, Part III*. Berlin, Heidelberg: Springer-Verlag, 2024, pp. 120–139.

[17] H. Muelder and T. Filatova. “One Theory - Many Formalizations: Testing Different Code Implementations of the Theory of Planned Behaviour in Energy Agent-Based Models”. In: *Journal of Artificial Societies and Social Simulation* 21.4 (2018), p. 5.

[18] J. Nellist. “Separation Anxiety in Horses”. In: *Animal Behaviour and Welfare Cases* (2023).

[19] C. Nicol. “Understanding equine stereotypies”. In: *Equine Vet J Suppl* 28 (1999), pp. 20–25.

[20] A. Notsu, H. Ichihashi, and K. Honda. “State and action space segmentation algorithm in Q-learning”. In: *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*. 2008, pp. 2384–2389.

[21] C. Rogsch. “The Influence of Moore and von-Neumann Neighbourhood on the Dynamics of Pedestrian Movement”. In: *Traffic and Granular Flow '15*. Ed. by V. L. Knoop and W. Daamen. Cham: Springer International Publishing, 2016, pp. 129–136.

[22] L. J. Rørvang MV Christensen JW and M. A. “Social Learning in Horses—Fact or Fiction?” In: *Frontiers in Veterinary Science* 5.212 (2018).

[23] K. Schneider. “What Is Learning?” In: vol. 15. 2024, pp. 779–799.

[24] A. Subramanian, S. Chitlangia, and V. Baths. “Reinforcement learning and its connections with neuroscience and psychology”. In: *Neural Networks* 145 (2022), pp. 271–287.

[25] V. Vinken et al. “Models of conditioned reinforcement and abnormal behaviour in captive animals”. In: *Behavioural Processes* 210 (2023), p. 104893.

[26] M. Wertheimer. “Obstacles to the integration of competing theories in psychology”. In: *Philosophical Psychology* 1.1 (1988), pp. 131–137.