International Journal of Astrobiology (2024), **23**, e21, 1–9
doi:10.1017/S1473550424000144

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Identifying optimal decision-making strategies and determining effective messaging to maximize the expected outcomes of potential human–extraterrestrial encounters

Ilan Fischer and Shacked Avrashi

School of Psychological Sciences, University of Haifa, Haifa, Israel
Corresponding author: Ilan Fischer; Email: ifischer@psy.haifa.ac.il

Received: 26 November 2023; **Revised:** 20 May 2024; **Accepted:** 26 July 2024

Keywords: behavioural game-theory, conflict management, cooperation, extraterrestrials, METI, SETI, similarity, strategic decision-making

Abstract
The question of whether extraterrestrials exist has driven both the search for extraterrestrial intelligence (SETI) and some attempts of messaging to extraterrestrial intelligence (METI). Nevertheless, no data-driven or theory-based behavioural policy has been suggested. Here we simulate a comprehensive set of human–extraterrestrial strategic interactions, modelled as two-by-two game-theoretic matrices. We examine a sample of possible outcomes by relying on the theory of subjective expected relative similarity (SERS), which takes into account both the expected payoffs and the extent of strategic similarity – the prospects of the opponent making identical choices. Simulation results suggest: focusing messaging efforts on signalling of complete strategic similarity, monitoring potential alien communications for similarity-indicating signals, and using risk-averse decision rules for policy planning and decision-making. The discussion puts forward three guidelines for METI initiatives and addresses the relevance of the findings to human conflict management.

Contents

Introduction

An interaction with extraterrestrial intelligence is a scenario with vague or completely unknown probabilities. If realized, it may turn out to be the most critical event in human history. While the search for extraterrestrial signals does not carry direct risks, human-sent messages may have unexpected consequences. Messages may be intercepted by civilizations with benevolent intensions, but they may also be intercepted by hostile civilizations and endanger the very existence of humanity (Dominik and Zarnecki, 2011; de Vladar, 2013; Todd and Miller, 2018). Renowned astrophysicist Stephen Hawking illustrated these risks by suggesting the possibility that an advanced alien civilization would have no problem wiping out the human race as easily as humans might wipe out a colony of ants (Cofield, 2015). In a similar manner, the science-fiction novel *The Dark Forest* suggests that the safest option for any species is to annihilate other life forms before they have the opportunity to do the same (Liu, 2008). Clearly, both scenarios suggest humans should avoid sending messages. However, some messaging attempts have already been carried out. For example, the *cosmic call* and the *lone signal* projects have sent radio messages to nearby stars (Braastad and Zaitsev, 2003; Gohring, 2013), the Apollo 11 spacecraft was fitted with a steel plaque stating 'Here men from the planet Earth first set foot upon the Moon … We came in peace for all mankind' (Johnson, 2008), and the Voyager space probes carry golden records that provide information about human nature, scientific and cultural achievements (NASA, n.d.).

Since contacting extraterrestrials may indeed carry serious hazards, messaging to extraterrestrial intelligence (METI) initiatives should be grounded in clear theory-driven models that maximize human outcomes and minimize potential risks. To this end, we apply game-theoretic modelling and simulate an exhaustive set of possible human–extraterrestrial interactions. More specifically, we rely on the theory of subjective expected relative similarity (SERS), which identifies choices that maximize expected values, not only by considering payoffs, but also by relying on *strategic similarity with the opponent* (Fischer, 2009, 2012; Fischer *et al.*, 2022).

In the following sections, we briefly explain the structure of two-by-two games by referring to the well-known Prisoner's Dilemma (PD) game, describe the SERS theory and its application to the PD game and proceed by extending the SERS solution to all two-by-two games. We then seed computer simulations with an exhaustive set of games, allowing modelling various possible human–extraterrestrial interactions. Examining simulation results, we define effective METI and SETI (search for extraterrestrial intelligence) theory-driven objectives, and consider humans' optimal decision-making strategies.

The PD game and SERS-driven choices

The PD game is a well-known model of cooperation and confrontation (Flood, 1958; Rapoport and Chammah, 1965; Axelrod, 1984). It is described by a two-by-two payoff matrix where each of two players may choose either a cooperative or a confrontational alternative (Fig. 1(a)). If both parties choose to cooperate, each player obtains the Reward (R) payoff; but if both choose to confront, each player obtains the Punishment (P) payoff. However, if one of the players chooses to confront while the other chooses to cooperate, the former obtains the Temptation (T) payoff and the latter obtains the Sucker's (S) payoff, where $T > R > P > S$. As derived from these inequalities, confrontation is a dominant strategy that allows players to protect themselves from exploitation (because $P > S$) and still retain the option to exploit a trusting opponent (because $T > R$). Many other decision rules, such as MaxiMin, MaxiMax (Rapoport, 1967) or the joint perspective expressed by the Nash equilibrium (Nash, 1950; Luce and Raiffa, 1957) concur with this conclusion, recommending PD players to always confront their opponents. In contrast to these principles, behavioural choices made by human participants show that while some players choose to confront, many others choose to cooperate (Rapoport and Chammah, 1965; Sally, 1995; Bornstein, 2003, 2004), suggesting that human behaviour is driven by other principles.

Bringing together theoretical rationality and observed behaviours, the SERS theory considers both the payoff values of the game and the extent of subjectively perceived *strategic similarity* with the opponent – the predicted prospects of both parties choosing identical alternatives in the forthcoming interaction, as subjectively experienced by each individual (Fischer, 2009, 2012; Fischer *et al.*, 2022; Fischer and Savranevski, 2023). Using subjective estimates of strategic similarity, denoted by p_s, and its complementary value, $1 - p_s$ (indicating the prospects of the opponent making dissimilar choices), players may, explicitly or rather implicitly, compute the expected values (EVs) for both possible choices of cooperation and confrontation. The EV of cooperation is given by $\mathrm{EV}_{\mathrm{cooperation}} = R \times p_s + S \times (1 - p_s)$, and the EV of confrontation is given by $\mathrm{EV}_{\mathrm{confrontation}} = P \times p_s + T \times (1 - p_s)$. Comparing $\mathrm{EV}_{\mathrm{cooperation}}$ with $\mathrm{EV}_{\mathrm{confrontation}}$ allows choosing the alternative that provides a higher EV. The derived decision rule may be expressed as: cooperate if $R \times p_s + S \times (1 - p_s) > P \times p_s + T \times (1 - p_s)$, and confront if $R \times p_s + S \times (1 - p_s) < P \times p_s + T \times (1 - p_s)$.

(a)

	Cooperate	Confront
Cooperate	R , R	S , T
Confront	T , S	P , P

(b)

	Column alternative - A	Column alternative - B
Row alternative - A	$V(AA)_{row}$, $V(AA)_{column}$	$V(AB)_{row}$, $V(AB)_{column}$
Row alternative - B	$V(BA)_{row}$, $V(BA)_{column}$	$V(BB)_{row}$, $V(BB)_{column}$

Fig. 1. *Panel a: the Prisoner's Dilemma game. Left and right payoffs in each cell indicate the respective payoffs for the row and column players. The game is defined by the inequalities:* T > R > P > S *(and in some experiments also requires* 2R > S + T*; Flood, 1958; Rapoport and Chammah, 1965). Panel b: a generic matrix structure showing two alternatives for the row and column players and their corresponding payoff values – V(row and column choices).*

Separating the payoffs of the interaction (i.e. T, R, S, P) from opponent's strategic similarity perception, p_s, SERS's decision rule may be rewritten as: cooperate if $p_s > (T-S)/(T-S+R-P)$, and confront the opponent if $p_s < (T-S)/(T-S+R-P)$. Further defining the *critical similarity threshold*, $p_s{*}$, which marks the switching point between cooperation and confrontation by $p_s{*} = (T-S)/(T-S+R-P)$, we obtain an abridged decision rule that recommends cooperation whenever $p_s > p_s{*}$, confrontation whenever $p_s < p_s{*}$ and being indifferent if $p_s = p_s{*}$.

SERS's solution has been shown to correctly predict human choices in critical single-shot games, including PD, Chicken and Battle of the Sexes (Fischer, 2009, 2012; Fischer and Savranevski, 2023). It has also been developed into an elaborate algorithm that outperforms many powerful algorithms in evolutionary computer simulations (Fischer *et al*., 2013).

Extending SERS to account for all two-by-two games

Examining SERS across all two-by-two games provides a comprehensive description of SERS's strategic recommendations. More specifically, it allows distinguishing between three types of games: (i) *similarity sensitive* games, for which the optimal choice depends, not only on the payoffs, but also on the perceived level of similarity with the opponent, p_s, (as shown for the PD game). In other words, one of the alternatives should be preferred if $p_s < p_s{*}$, while the other alternative should be preferred if $p_s > p_s{*}$. The level of $p_s{*}$ is dependent only on the payoffs of each specific game (and may therefore differ between games). The level of perceived strategic similarity, p_s, is subjectively estimated by each player, expressing the prospects of the opponent making identical choices (and may therefore differ between players). Importantly for the hereby considered humans–extraterrestrials interaction, a player expecting to play a similarity sensitive game may attempt influencing the choice of the opponent by messaging indications of strategic similarity that motivate opponent's cooperation. (ii) Games that are *non-similarity-sensitive* where SERS's decision rule recommends choosing the same alternative regardless of the estimated level of p_s. (iii) Games that are similarity sensitive for one of the players (either the row or the column player) but not for the other.

To quickly identify whether games are similarity sensitive (independently for each player), one may compare the preferred choice made under the assumption of $p_s = 1$ (i.e. both players are assured to choose identical alternatives) and the preferred choice made under the assumption of $p_s = 0$ (i.e. both players are assured to choose different alternatives). If the preferred choices differ, the game is similarity sensitive from the perspective of the player whose payoffs are considered, and vice versa.

SERS computation requires identifying the cells that constitute players' similar alternatives. Therefore, it is helpful to standardize all games by arranging the cells of each matrix in a fixed order that assures similar and dissimilar alternatives are located in designated respective cells, allowing to correctly associate similar-alternative cells with p_s, and dissimilar-alternative cells with $1 - p_s$. Notice that switching rows, columns or both rows and columns does not change the game but only its presentation. Games that are *not fully symmetric* do not offer *identical* payoffs to both players.

https://doi.org/10.1017/S1473550424000144 Published online by Cambridge University Press

For these games, the similarity of the alternatives is not reflected by identical payoffs of each player, but by the relative proximity of the payoffs. In other words, by choosing similar alternatives in asymmetric games, players obtain payoffs that are not identical, but are more similar to each other than the payoffs obtained by choosing dissimilar alternatives.

Figure 1(b) shows a generic two-by-two game structure, with row and column players' alternatives denoted by A and B, where AA and BB represent similar alternatives. The payoff values for row and column players are denoted by V_{row}, and V_{column}, respectively. Row player's payoffs are defined as: $V(\text{AA})_{\text{row}}$, $V(\text{AB})_{\text{row}}$, $V(\text{BA})_{\text{row}}$, $V(\text{BB})_{\text{row}}$; and column's player payoffs are defined as: $V(\text{AA})_{\text{column}}$, $V(\text{AB})_{\text{column}}$, $V(\text{BA})_{\text{column}}$, $V(\text{BB})_{\text{column}}$.

After correctly aligning the cells in each matrix, we rewrite SERS's decision rule (for the row player) as:

Choose A if $p_s \times V(\text{AA}) + (1 - p_s) \times V(\text{AB}) > p_s \times V(\text{BB}) + (1 - p_s) \times V(\text{BA})$, choose B if $p_s \times V(\text{AA}) + (1 - p_s) \times V(\text{AB}) < p_s \times V(\text{BB}) + (1 - p_s) \times V(\text{BA})$, otherwise be indifferent. Further defining $p_s^* = [V(\text{BA}) - V(\text{AB})]/[V(\text{BA}) - V(\text{AB}) + V(\text{AA}) - V(\text{BB})]$ generates the abridged decision rule: choose A if $p_s > p_s^*$ and choose B if $p_s < p_s^*$ (Fig. 1(b)). Note that deriving the parallel decision rule for the column player requires switching $V(\text{AB})$ and $V(\text{BA})$.

Notably, players who expect their opponent to play a similarity sensitive game may try to influence the way they are being perceived by the opponent. They may attempt to convey a higher extent of strategic similarity, consequently expecting the opponent to choose a more cooperative alternative. If opponent's personal and social properties (e.g. beliefs, goals, traits, history or culture) are known, one may emphasize those properties that are more similar to one's own properties, and use them as indicators or proxies of strategic similarity. If no information is available, as in the case of unknown extraterrestrial entities, strategic similarity signals may still be conveyed by relying on rather fundamental concepts and universal properties (see Discussion).

Figure 2 illustrates three games played by two parties, denoted as humans and extraterrestrials, labelled in accord with the property of the game being similarity sensitive for both, one or none of the parties. Note that choosing an SERS-based alternative does not require calculating the games' similarity threshold or determining whether the game is similarity sensitive. In practice, each player only needs to correctly associate the probability of strategic similarity, p_s, with the corresponding payoffs (and the complimentary probability $1 - p_s$ with its corresponding payoffs) and compute the EV for each of the two alternatives. Finally, the player should compare both EVs and choose the alternative

(a) Extraterrestrials

Humans	α	β
A	10 , 10	0 , 1
B	1 , 0	3 , 3

Similarity-sensitive

$$EV(A) = 10p_s + 0(1 - p_s)$$
$$EV(B) = 1(1 - p_s) + 3p_s$$

(b) Extraterrestrials

Humans	α	β
A	10 , 10	3 , 0
B	0 , 3	5 , 5

Non-similarity-sensitive

$$EV(A) = 10p_s + 3(1 - p_s)$$
$$EV(B) = 0(1 - p_s) + 5p_s$$

(c) Extraterrestrials

Humans	α	β
A	10 , 8	0 , 1
B	3 , 4	5 , 5

One player similarity-sensitive

$$EV(A) = 10p_s + 0(1 - p_s)$$
$$EV(B) = 3(1 - p_s) + 5p_s$$

Fig. 2. *Examples of two-by-two game matrices, here described as games taking place between humans and extraterrestrials. Humans choose between alternatives A and B, and extraterrestrials choose between alternatives α and β. Each cell shows the payoff values obtained by humans (left) and by extraterrestrials (right). EV(A) and EV(B) denote the expected value for the* human *party when choosing alternatives A or B, and* p_s *denotes the probability the* human *party assigns to the prospects of both parties making identical choices (i.e. either Aα or Bβ). Matrix* ***a*** *shows a game that is* similarity-sensitive *for both parties. Matrix* ***b*** *shows a* non-similarity sensitive *game. Matrix* ***c*** *shows a game that is* similarity-sensitive *only for the human party.*

https://doi.org/10.1017/S1473550424000144 Published online by Cambridge University Press

that provides the higher one. As shown in previous studies, actual human choices are highly correlated with this computational model, meaning that human individuals naturally and intuitively (without engaging in formal computations) associate similarity perceptions with the corresponding payoffs (Fischer, 2009, 2012; Fischer and Savranevski, 2023).

Simulations

To simulate a representative and comprehensive set of possible two-by-two interactions, we randomly sample payoffs from a quasi-continuous and equiprobable range (0, 1, …, 99, 100) to generate 10^5 games, representing an illustrative and sufficiently large sample of games. We then distinguish between strategic similarity expectations of the parties, each assuming 11 levels (p_s = 0, 0.1, …, 0.9, 1), and compute SERS-based EVs of the human party for all 10^5 games × 11 levels of human similarity expectations × 11 levels of extraterrestrial similarity expectations. Before computing the SERS-based EVs (Fig. 2), the simulation software examines each game and properly aligns the alternatives of both parties by determining which human alternative best corresponds to which extraterrestrial alternative, allowing to correctly associate p_s and $1 - p_s$ probabilities with similar and dissimilar choices. The alignment may be tested by applying a correlation-based or a difference-based procedure. Both methods compare the two possible configurations of the alternatives and test the extent to which the players face the same strategic problem under each of both possibilities, thus pointing to the more similar and to the more dissimilar choices. Notably, both configurations represent the same game; therefore the realignment does not change the strategic nature of the encounter (see Supplementary materials). Since both the correlation-based and the difference-based alignment procedures generate rather identical outcomes, the hereby reported results are derived from the correlation-based alignment method (the parallel results computed using the difference-based alignment are presented in the Supplementary materials). After alignment, the software computes the EVs for each alternative and selects the alternative with the higher EV, separately for humans and for extraterrestrials, thus determining both parties' preferred choices. Finally, the actual payoffs in the chosen cells (determined by the intersection of the selected alternatives) are averaged for all different combinations of human and extraterrestrial strategic similarity perceptions. The entire process is illustrated in Fig. 3.

Figure 4 depicts simulation results showing the expected human payoffs across 11 possible *human-held* similarity perceptions (p_s = 0, 0.1, 0.2, …, 0.9, 1), each averaged across 11 possible levels of *extraterrestrial* similarity perceptions (p_s = 0, 0.1, 0.2, …, 0.9, 1), for a total of 1.21×10^7 games. Panels a and b present separate results for similarity sensitive games, non-similarity-sensitive games, human-only similarity-sensitive games and extraterrestrial-only similarity-sensitive games (each game category comprises approximately 25% of the simulated sample); both panels also show the average payoffs across all game types. Panel a shows that the maximal *human* mean payoff is obtained when making decisions under the assumption of $p_s = 0.5$. Panel b shows that the maximal *human* mean payoff is obtained when extraterrestrials expect a maximal strategic similarity with humans ($p_s = 1$). Panel c depicts a two-dimensional heat map that shows human mean payoffs for all tested combinations of human and extraterrestrial similarity expectations, revealing a clear strategic sweet spot at the {0.5, 1} coordinate, thus indicating that the maximal mean human payoff is obtained when humans make decisions while assuming 50% strategic similarity with extraterrestrials and extraterrestrials make decisions while assuming 100% strategic similarity with humans. Moreover, the {0.5, 1} coordinate provides not only an optimal average expected payoff, but also the best possible payoff for 73% of all simulated games (i.e. no other coordinates provide better payoffs), as well as the lowest variance of human expected payoffs across all games (SD = 24.6; see simulation results in the Supplementary materials). These results suggest that as long as no strategic similarity level with extraterrestrials can be specified, *risk-neutral payoff-maximizing* humans should: (i) make decisions while assuming $p_s = 0.5$, and (ii) direct all messaging attempts to convey a strategic similarity level of $p_s = 1$.

https://doi.org/10.1017/S1473550424000144 Published online by Cambridge University Press

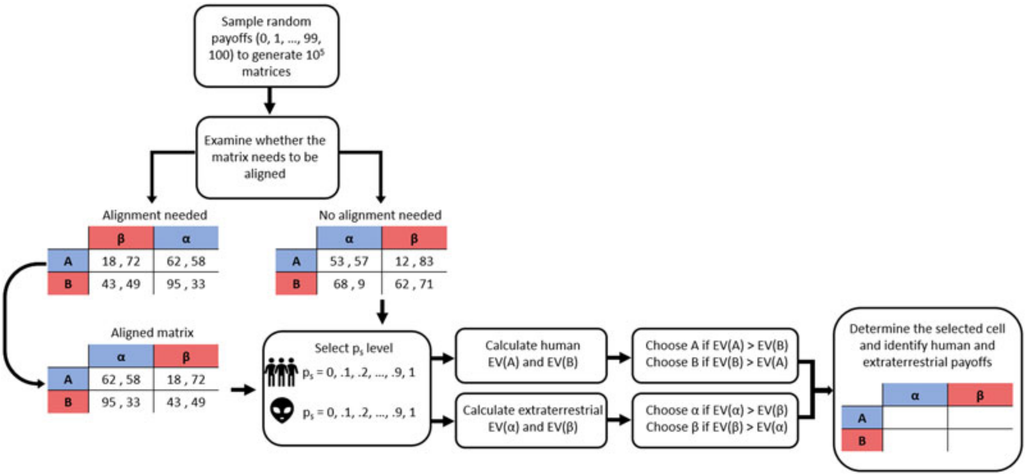

Fig. 3. *Illustration of the simulation process. Random payoff values are sampled to generate 10^5 matrices. Each matrix is then separately tested. If necessary, an alignment protocol is activated according to the correlation-based method (see also the difference-based method in the Supplementary materials). Each matrix is associated with all tested* p_s *levels (11 strategic similarity perceptions for humans, and 11 perceptions for extraterrestrials), allowing calculating SERS-based EV(A), EV(B), EV(α) and EV(β). The alternatives with the higher EVs are separately selected for humans and for extraterrestrials, allowing to determine the jointly chosen cell and to assign its payoffs to the respective parties.*

Discussion

Simulation results show that under the conditions of strict uncertainty, where the probabilities of strategic similarity with extraterrestrials cannot be specified, humans obtain their maximal payoffs when extraterrestrials perceive a maximal strategic similarity with humans ($p_s = 1$) and humans assume a 50% strategic similarity with extraterrestrials ($p_s = 0.5$). Together these results generate an optimal outcome at the {0.5, 1} coordinate, as depicted in Fig. 4(c). This theoretical result sets a clear METI objective – messaging similarity indicating signals that motivate extraterrestrials' cooperation and enhance human expected payoffs. It also highlights the risks associated with wrong messages that are likely to indicate differences (Todd and Miller, 2018) rather than similarities. To mitigate these risks, we propose the following three simple rules:

(1) Avoid any content that conveys human goals, aspirations or strategic intensions, whether they express cooperative or confrontational attitudes, as such content may indicate the existence of strategic differences.
(2) Avoid any content that reflects human biological properties, historical events or cultural and scientific achievements, as this human-centric information may also indicate differences rather than similarities.
(3) Provide content-free and neutral similarity-indicating cues (see Fig. 5 for examples). Since a neutral stimulus may carry unknown meanings in another culture, messages should comprise an assortment of unrelated stimuli, thus minimizing the prospects of repeating an unintended meaning.

Clearly, we must also consider the possibility that extraterrestrials have reached similar conclusions and may have already broadcasted similarity indicating messages. If detected, such messages must be carefully examined, bearing in mind the potential for genuinely cooperative intentions, as well as the possibility of strategic motivations aimed at increasing extraterrestrials' payoffs in a prospective interaction.

https://doi.org/10.1017/S1473550424000144 Published online by Cambridge University Press

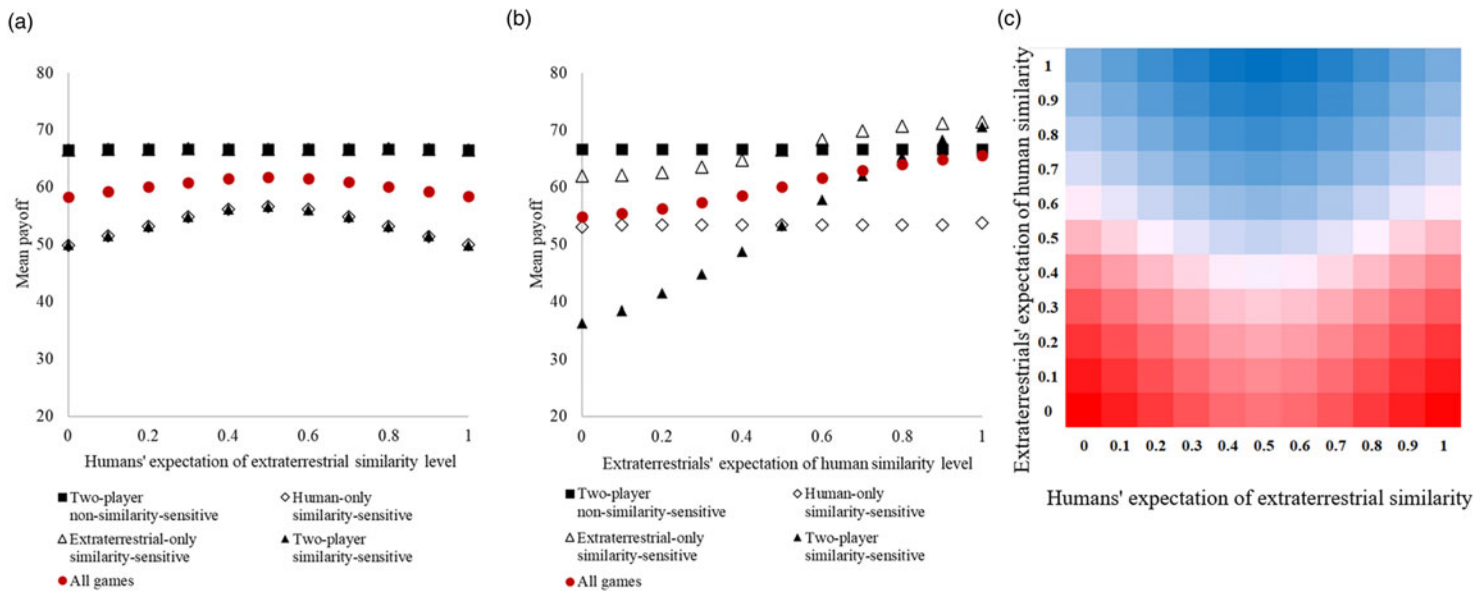

Fig. 4. *Simulation results. Panel a depicts mean human payoffs across various human expectations of extraterrestrial strategic similarity levels, calculated separately for each game type and across all games. Panel b depicts mean human payoffs across various extraterrestrial expectations of human strategic similarity levels, calculated separately for each game type and across all games. Panel c depicts a two-dimensional heat map of human mean payoffs calculated across 11 human × 11 extraterrestrial probabilities of similarity. Low and high human payoffs are denoted by shades of red and blue, respectively.*

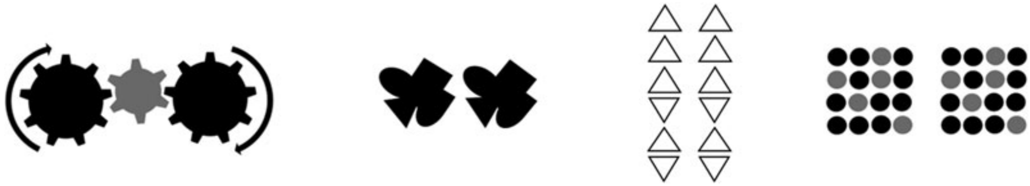

Fig. 5. An assortment of plain, neutral and unrelated visual images, each indicating complete similarity.

The optimal outcome associated with humans applying $p_s = 0.5$, derived from averaging across all possible similarity perceptions held by extraterrestrials, is a risk-neutral result (Weber and Milliman, 1997). It does neither reflect the hazards associated with overestimation (i.e. actual $p_s < 0.5$), nor the potential harms associated with underestimation (i.e. actual $p_s > 0.5$). Since the risks may seem more crucial for humanity than the potential gains, we propose making decisions based on risk-averse or pessimistic decision rules, such as MaxiMin (Luce and Raiffa, 1957) or MiniMax Regret (Loomes and Sugden, 1982), to avoid the worst-case scenario and minimize future regret. Nonetheless, even this cautious approach does not guarantee the avoidance of cataclysmic consequences.

Three additional comments are warranted. First, while the described results are derived from modelling human–extraterrestrial interactions using two-by-two games, transitioning to more complex and ecologically-valid *m*-by-*n* games is likely to alter similarity thresholds. Moreover, games with more than two alternatives allow for defining multiple similarity thresholds, each providing a decision criterion for preferring one alternative over another (or over a set of other alternatives). Nevertheless, perceiving high levels of strategic similarity motivates choosing cooperative alternatives even in these more complex games. Other non-game-theoretic research paradigms also support SERS's prediction, showing that human participants associate similarity with interpersonal attraction (Byrne, 1971), affiliation (Lakin and Chartrand, 2003), rapport (Chartrand and Bargh, 1999) and generosity (van Baaren *et al*., 2003). This converging evidence suggests that humans (though not necessarily extraterrestrials) associate similarity not only with the choice of cooperative alternatives but also with various forms of benevolent behaviour.

Second, since new data may become available through scientific research or as a result of actual encounters (assuming humanity survives the consequences), the recommendations derived from the present simulations will need to be updated and replaced by new and more resolved information. Therefore, scientists and policymakers should not only prepare for a first encounter, but continually monitor new evidence, plan ahead and update various applicable policies.

Third, signalling strategic similarity is advantageous not only in human–extraterrestrial interactions but also in preventing and resolving human conflicts. Unlike unknown entities, human populations exhibit historical, cultural and behavioural properties. Some of these properties may indicate dissimilarity and should therefore be played down, while others may reveal similarity and should be utilized to promote cooperation.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S1473550424000144

Competing interests. None.

References

Axelrod RM (1984) *The Evolution of Cooperation*. New York, NY: Basic Books.

Bornstein G (2003) Intergroup conflict: individual, group, and collective interests. *Personality and Social Psychology Review* **7**, 129–145.

Bornstein G (2004) Cooperation in intergroup and single-group social dilemmas. In Suleiman R, Fischer I, Budescu D and Messick D (Eds), *Contemporary Psychological Research on Social Dilemmas*. Cambridge University Press, pp. 227–247.

Braastad R and Zaitsev A (2003) *Synthesis and Transmission of Cosmic Call* 2003 *Interstellar Radio Message*. Moscow: Kotelnikov Institute of Radio-engineering and Electronics. Available at http://www.cplire.ru/html/ra&sr/irm/CosmicCall-2003/
Byrne DE (1971) *The Attraction Paradigm*. Academic Press.
Chartrand TL and Bargh JA (1999) The chameleon effect: the perception–behavior link and social interaction. *Journal of Personality and Social Psychology* **76**, 893–910.
Cofield C (2015) *Stephen Hawking: Intelligent Aliens Could Destroy Humanity, But Let's Search Anyway Space*. New York, NY: space.com. Available at https://www.space.com/29999-stephen-hawking-intelligent-alien-life-danger.html
de Vladar HP (2013) The game of active search for extra-terrestrial intelligence: breaking the 'great silence'. *International Journal of Astrobiology* **12**, 53–62.
Dominik M and Zarnecki JC (2011) The detection of extra-terrestrial life and the consequences for science and society. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **369**, 499–507.
Fischer I (2009) Friend or foe: subjective expected relative similarity as a determinant of cooperation. *Journal of Experimental Psychology: General* **138**, 341–350.
Fischer I (2012) Similarity or reciprocity? On the determinants of cooperation in similarity-sensitive games. *Psychological Inquiry* **23**, 48–54.
Fischer I and Savranevski L (2023) The effect of similarity perceptions on human cooperation and confrontation. *Scientific Reports* **13**, 1–10.
Fischer I, Frid A, Goerg SJ, Levin SA, Rubenstein DI and Selten R (2013) Fusing enacted and expected mimicry generates a winning strategy that promotes the evolution of cooperation. *National Academy of Sciences* **110**, 10229–10233.
Fischer I, Levin SA, Rubenstein DI, Avrashi S, Givon L and Oz T (2022) Interacting with others while reacting to the environment. *Behavioral and Brain Sciences* **45**, 27–28.
Flood M (1958) Some experimental games. *Management Science* **5**, 5–26.
Gohring N (2013) *Lone Signal Aims to Send 'Hello!' Tweets to Extraterrestrials*. Atlanta, GA: CNN. Available at https://money.cnn.com/2013/06/17/technology/enterprise/lone-signal/index.html
Johnson S (2008) Red, white, & blue. U.S. flag at home on the Moon. *Houston History Magazine* **6**, 60–63.
Lakin JL and Chartrand TL (2003) Using nonconscious behavioral mimicry to create affiliation and rapport. *Psychological Science* **14**, 334–339.
Liu C (2008) *The Dark Forest*. New York, NY: Tor Books.
Loomes G and Sugden R (1982) Regret theory: an alternative theory of rational choice under uncertainty. *The Economic Journal* **92**, 805.
Luce RD and Raiffa H (1957) *Games and Decisions: Introduction and Critical Survey*. New York, NY: Wiley. Available at https://psycnet.apa.org/record/1958-00040-000
NASA (n.d.) *Voyager – The Golden Record*. Retrieved 7 September 2023, from https://voyager.jpl.nasa.gov/golden-record/
Nash JF (1950) Equilibrium points in n -person games. *Proceedings of the National Academy of Sciences* **36**, 48–49.
Rapoport A (1967) Exploiter, leader, hero, and martyr: the four archetypes of the 2×2 game. *Behavioral Science* **12**, 81–84.
Rapoport A and Chammah AM (1965) *Prisoner's Dilemma: A Study in Conflict and Cooperation*. Ann Arbor, MI: University of Michigan Press.
Sally D (1995) Conversation and cooperation in social dilemmas. *Rationality and Society* **7**, 58–92.
Todd PM and Miller GF (2018) The evolutionary psychology of extraterrestrial intelligence: are there universal adaptations in search, aversion, and signaling? *Biological Theory* **13**, 131–141.
van Baaren RB, Holland RW, Steenaert B and van Knippenberg A (2003) Mimicry for money: behavioral consequences of imitation. *Journal of Experimental Social Psychology* **39**, 393–398.
Weber EU and Milliman RA (1997) Perceived risk attitudes: relating risk perception to risky choice. *Management Science* **43**, 123–144.

https://doi.org/10.1017/S1473550424000144 Published online by Cambridge University Press