

Articles

Nick and Ibrahim's Prisoner's Dilemma in Split or Steal: Relevance and Goal Conciliation in Golden Balls

O Dilema do Prisioneiro de Nick e Ibrahim em Split or Steal: Relevância e Conciliação de Metas em Golden Balls

Fábio José Rauen*
Andressa Bregalda Belan**

ABSTRACT

We examine in this article the final round of an episode of the British game show Golden Balls, in which Nick and Ibrahim negotiate the £13,600 jackpot. In Split or Steal, both players lose the prize if they both choose to steal, and they split the prize if they both choose to split. However, if one player chooses to steal while the other chooses to split, the former takes the entire prize. In the episode, Nick proposes stealing the prize to split it later, and despite Ibrahim's initial skepticism due to the proposal not being sanctioned by the game's rules, he eventually agrees. Surprisingly, Nick also opts to split, indicating a strategic move to prevent Ibrahim from claiming the entire prize. Using Relevance and Goal Conciliation Theories, we examine how the negotiation context influences practical-epistemic vigilance strategies, suggesting advantages for proponents in stressful decision-making situations. We argue that the game conditions, akin to the Prisoner's Dilemma in Game Theory, restrain the detection of bluffs and cheats and the consideration of countermoves.

Keywords: Relevance Theory; Goal Conciliation Theory; Practical-Epistemic Vigilance; Prisoner's Dilemma; Golden Balls.

* Docente do Programa de Pós-graduação em Ciências da Linguagem da Universidade do Sul de Santa Catarina. Santa Catarina – Brasil. <https://orcid.org/0000-0002-1096-7253>. E-mail: fabio.rauen@gmail.com.

** Docente do Instituto Federal de Santa Catarina, Câmpus Tubarão. Santa Catarina – Brasil. <https://orcid.org/0000-0003-3196-4948>. E-mail: andressa.belan@ifsc.edu.br.

RESUMO

Examinamos neste artigo a rodada final de um episódio do game show britânico Golden Balls, no qual Nick e Ibrahim negociam o prêmio acumulado de £ 13.600. Em Split or Steal, ambos os jogadores perdem o prêmio se escolherem roubá-lo e o dividem se ambos optarem por dividi-lo; porém, aquele que decide roubá-lo enquanto o outro escolhe dividi-lo fica com todo o prêmio. No episódio, Nick propõe roubar o prêmio para dividi-lo depois do jogo, e Ibrahim, embora desconfiado dessa proposta não sancionada pelas regras do jogo, acaba por concordar. Surpreendentemente, Nick também escolhe dividir, o que sugere uma manobra para impedir Ibrahim de levar o prêmio inteiro. Utilizando as teorias de relevância e de conciliação de metas, exploramos como o contexto da negociação afeta estratégias de vigilância prático-epistêmica, sugerindo vantagens aos proponentes em circunstâncias estressantes de tomada de decisão. Argumentamos que as condições do jogo, semelhantes ao Dilema do Prisioneiro em teoria dos jogos, restringem a detecção de blefes e trapaças, bem como a consideração de contragolpes.

Palavras-chave: Teoria da relevância; Teoria de conciliação de metas; Vigilância prático-epistêmica; Dilema do Prisioneiro; Golden Balls.

1. Introduction

Golden Balls was a British game show aired on the ITV Network from June 18, 2007, to December 18, 2009, hosted by Jasper Carrott, produced by Endemol UK, and filmed at the BBC Television Centre (Wikipedia, n.d.). It had four rounds, in which contestants eliminated each other based on the cash and killer balls they had. In the last round—*Split or Steal*—, the two finalists had to decide in approximately 30 seconds whether to split or steal the jackpot, in a weaker version of the *Prisoner's Dilemma*.

The Prisoner's Dilemma is a widely recognized scenario in Game Theory.¹ It illustrates why rational agents may not cooperate even when it aligns with their best interests. In the classic example, police officers question

1. *Game theory* analyzes how people make strategic decisions when interacting with others. Inspired by von Neumann and Morgenstern's 1944 ideas, it uses mathematical methods to find solutions for different types of games, such as zero-sum games, where one player's gain is another player's loss, and non-zero-sum games, where players can have shared or conflicting interests. The theory examines cooperative and non-cooperative strategies in decision-making, where players may either coordinate their choices or act independently (Sierralta Ríos, 2017; Fiani, 2015; Bêrni & Fernandez, 2014). Strictly speaking, the *Prisoner's Dilemma* is a paradigm of social cooperation and conflict in which players must weigh self-interest against social responsibility.

two suspects of a crime separately: if neither confesses, the authorities imprison both for a few months for a lesser offense; if both confess, they face a five-year prison term; if only one confesses, that individual is released, while the other serves a ten-year sentence. Because coordinating outcomes is impossible, each prisoner has a dominant strategy to confess, regardless of the other's decision.² However, if both confess, they end up in a less favorable situation than if both remain silent, demonstrating that individual rationality does not always align with collective rationality.

In *Split or Steal*, the issue is slightly similar: if both contestants opt to share the jackpot, each contestant receives half of its value; if only one opts to steal, that individual claims the entire prize; if both choose to steal, they walk away without any winnings. Within this context, even though the optimal strategy is to share, this decision unfolds within a negotiation environment tinged with the suspicion that the opponent may not cooperate, even though the one willing to steal also risks losing everything.³ Given the evident influence of argumentation on decision outcomes, we argue that the integration of dynamic pragmatic-cognitive processes, as delineated in Relevance Theory and Goal Conciliation Theory, is imperative for comprehensive decision-making analyses. Therefore, analyses traditionally confined to mathematical modeling stand to gain substantial benefits from a more inclusive examination of cognitive-pragmatic issues.

Sperber and Wilson's (1986, 1995) *Relevance Theory* is a cognitive approach to Pragmatics based on the theoretical notion of relevance. Relevance is a potential property of cognitive inputs, considered so when their positive effects outweigh the processing effort. This occurs when they reinforce, contradict, or generate new implications through interaction with previous cognitive assumptions. The authors derived two principles of relevance from their analysis of human cognition and communication. The cognitive principle states that human cognition tends to be geared toward the maximization of relevance. The communicative principle states that every act of ostensive communication conveys a presumption of its optimal

2. The case is an example of *Nash Equilibrium* because no player can enhance their outcome by changing their strategy in the context of others' choices. It involves selecting a strategic move aligning with a player's goals when others do the same (Fiani, 2015).

3. Note that there is a critical difference in *Split or Steal* because, as Wegmüller (2020) points out, there is no penalty when both players choose to steal the prize. While in the classical version of the dilemma, both prisoners face a five-year prison term, in this weaker version, both contestants lose what they expected to gain.

relevance. This means that the utterances or other ostensive stimuli are at least relevant enough to be worth the hearer's processing effort and are the most relevant ones compatible with the speaker's abilities and preferences. The hearer uses this presumption to guide their comprehension procedure, which is a heuristic that involves inferring the speaker's intended meaning by following a path of least effort, considering hypotheses in order of accessibility, and stopping when their expectations of relevance are met (or abandoned).

Rauen's (2014) *Goal Conciliation Theory* is a cognitive approach to Pragmatics based on a four-stage model for intentional action planning. The model begins by projecting a goal and proceeds through formulating, executing, and evaluating one or more antefactual abductive hypotheses related to that goal. Antefactual abductive hypotheses are hypothetical scenarios that the agents design before acting and evaluate according to their expected relevance and feasibility. Goal conciliation occurs when a state of the world aligns with the agent's goals. It can be self-conciliation when the agent achieves their own goals, or hetero-conciliation, when the agent achieves their goals collaboratively through relevance-driven ostensive-inferential communication.

In the context of Relevance Theory, Sperber et al. (2010) claim that individuals engage in processes of epistemic vigilance to ensure that communicative interactions yield positive cognitive effects, considering the contingency of both accidental (misinformation) and intentional (deception and bluffing) disinformation. In the context of Goal Conciliation Theory, Luciano (2019) argues that this engagement can be more accurately characterized as practical-epistemic vigilance. According to her perspective, individuals not only gather advantageous knowledge but also scan practical intentions.

Building on these aspects, this article examines how Nick and Ibrahim negotiate the £13,600 jackpot in *Split or Steal*, assuming the hypothesis that this interaction is asymmetric.⁴ Accordingly, we argue that one participant proposes intentional action plans for the egotistic or altruistic hetero-conciliation of their practical goals. Simultaneously, the other participant infers these goals to guide decision-making, engaging mechanisms of practical-epistemic vigilance in the process

2. What kind of proposal is this?

After Jasper Carrott explains the rules and authorizes the negotiation, Nick surprises everyone by urging Ibrahim to trust him, assuring him that he will steal the prize and split it afterward.

- (1) Nick (02:38): Ibrahim, I want you to trust me.
- (2) Nick (02:40): One hundred percent, I am going to pick the steal ball.

Accordingly, our first step is to mobilize the relevance-theoretic descriptive-explanatory architecture to examine how Ibrahim (and the audience) process these two utterances. As we anticipate, individuals presume that information directed toward them is sufficiently rich in positive effects and economical enough in terms of effort to deserve attention. Thus, communication involves conveying information with optimal relevance, and the criterion that ensures the selection of at least one pertinent interpretation among competing alternatives is consistency with the presumption of optimal relevance, which governs the relevance-theoretic comprehension process.

In sum, this procedure will generate explicatures of the uttered sentences, considering the linguistic, logical, and encyclopedic inputs of each of their concepts. If these explicatures, combined with contextual assumptions, fail to yield an interpretation that satisfies the presumption of optimal relevance, the hearer may look for additional implicated premises to derive further implicated conclusions (implicatures). This process continues until the hearer arrives at a satisfactory propositional representation (if possible) of the speaker's intentions.

In general, the logical forms of uttered sentences are not fully propositional. In these cases, assigning a truth value to their propositional content involves developing their logical form until it becomes fully propositional. An explicature in these terms involves assigning encyclopedic entries to the constituents of the logical form, regardless of whether those constituents have explicit linguistic counterparts in the sentences.

Let us return to Nick's first utterance:

- (1) Nick: Ibrahim, I want you to trust me.

There are several ways to examine how Ibrahim (and the audience) would develop the logical form of the sentences within Nick's utterance

(Figure1). Here, we assume that the logical form would be organized around a superordinate sentence, stating that “someone_i [supposedly Nick_i] wants something for someone_j [supposedly Ibrahim_j] at some place and time,” and a subordinate sentence developing what Nick wants, namely, that “someone_j [supposedly Ibrahim_j] trusts someone_i [supposedly Nick_i].”

Figure 1 – Processing of Nick’s utterance (1)

Linguistic Form	I	want	Ø	you	to trust	me	Ø	Ø
Logical Form	someone _i	wants	for someone _j	something		somewhere		sometime
				Someone _j to trust someone _i				
Conceptual Form	NICK	WANTS	FOR IBRAHIM	IBRAHIM	TO	NICK	AT THE SPLIT OR STEAL NEGOTIATION’S PLACE	AT THE SPLIT OR STEAL NEGOTIATION’S TIME

Source: Figure developed by the authors.

Therefore, we can represent the explicature of Nick’s first utterance by a cognitive assumption, such as A₁ below, which is both a fully propositional and an accurate description of what Nick wants.

A₁ – Nick wants Ibrahim to trust Nick at the *Split or Steal* negotiation’s place and time (explicature of Nick’s first utterance).

Furthermore, we can add the speech act involved in the utterance, because it is also true that Nick expresses his desire for Ibrahim’s trust.

A₁ – Nick says that Nick wants Ibrahim to trust Nick at the *Split or Steal* negotiation’s place and time (higher level explicature of Nick’s first utterance, including the speech act).

We can analyze Nick’s second utterance in similar terms:

(2) Nick: One hundred percent, I am going to pick the steal ball.

A₂ – Nick says that Nick is one hundred percent going to pick the steal ball at the final of the *Split or Steal* negotiation’s place and time (higher level explicature of Nick’s second utterance).

However, these are probably not Nick’s intended interpretations. We argue that Nick wants Ibrahim to include assumptions A₁-A₂ above into the

context of *Golden Balls* rules and provides him with two unfavorable options because Ibrahim loses his £6,800 prize in both alternatives.

Ibrahim's first option – If Ibrahim chooses to split the £13,600 prize, and Nick chooses to steal the £13,600 prize, then Ibrahim loses the £6,800 prize, and Nick wins the £13,600 prize (implied premise by *conjunctive modus ponens*—formally, $A_3 \wedge A_4 \rightarrow A_5 \wedge A_6$ —derived from the encyclopedic knowledge of the game rules).⁵

Ibrahim's second option – If Ibrahim chooses to steal the £13,600 prize, and Nick chooses to steal the £13,600 prize, then Ibrahim loses the £6,800 prize, and Nick loses the £6,800 prize (implied premise by *conjunctive modus ponens*—formally, $A_7 \wedge A_4 \rightarrow A_5 \wedge A_8$ —derived from the encyclopedic knowledge of the game rules).

In other words, Nick's proposal takes the following form: if Ibrahim trusts Nick's word that he will steal the prize ($A_1 \wedge A_2$), then either Ibrahim splits the prize, and Nick takes the entire prize home ($A_3 \wedge A_4 \rightarrow A_5 \wedge A_6$), or Ibrahim steals the prize, and both go home with nothing ($A_7 \wedge A_4 \rightarrow A_5 \wedge A_8$).

Formally:

Nick's proposal: $(A_1 \wedge A_2) \rightarrow (A_3 \wedge A_4 \rightarrow A_5 \wedge A_6) \vee (A_7 \wedge A_4 \rightarrow A_5 \wedge A_8)$.

According to Wilson (2004), it is naive to presume the speaker never thwarts hearers' expectations of relevance. Furthermore, it is unrealistic to assume speakers are immune to mistakes or incapable of deceptive purposes. While speakers may inaccurately predict what hearers will consider optimally relevant or misjudge the truthfulness of information—thereby conveying false data even in good faith—it is crucial to recognize that speakers do not always act benevolently and may deliberately disseminate disinformation. Thus, we can consider degrees of sophistication in interpretation, such that the hearers may accept an interpretative hypothesis either because it is relevant in the expected manner (naive interpretation), because speakers design it as relevant (cautious interpretation), or because speakers design

5. Relevance Theory assumes there is a deductive mechanism deriving the intended meaning of an utterance from its logical form and contextual information. For example, the *elimination-and* rule states that if a conjunction of two propositions is true, each proposition is true separately. The *modus ponens* rule states that if the antecedent of a conditional proposition is true, then its consequent is true. If the antecedent of a conditional proposition contains a conjunction of true propositions, we can first eliminate each conjunct individually and then apply the *modus ponens* rule.

it for the hearers to see as relevant (sophisticated interpretation) (Wilson, 2004). Therefore, it is worth conceiving not only a concept of epistemic vigilance—according to Sperber et al. (2010), toward the sources of information (vigilance against cheaters) or the information itself (vigilance against deceptions and tricks)—but also a concept of practical vigilance—according to Luciano (2019), toward people’s interests.

As we anticipate, Rauen (2014) models human proactivity in four stages. Thus, in stage [1], the individuals project a future state of goal Q ; in stage [2], they abduce at least one antefactual hypothesis PQ (intentional action plan) positing an antecedent action P that enables the achievement of this goal Q ; in stage [3], they execute the antecedent action P with the expectation of achieving this goal Q ; and in stage [4], they deductively check the achievement Q' in the context of their antecedent action P .⁶

In this model, the first three stages are abductive. The reasoning starts with a future-projected goal Q in stage [1] and the emergence of at least one antefactual abductive hypothesis PQ for achieving that goal Q in stage [2], leading to the choice or decision to execute the antecedent action P in stage [3]. The last three stages, on the other hand, are deductive. The reasoning starts with the abductive antefactual hypothesis PQ in stage [2]—taken as a major premise—and the antecedent action P in stage [3]—taken as a minor premise—leading to the conclusion Q' in stage [4] by *modus ponens*.

We can see this abductive-deductive architecture in Figure 2.

Figure 2 – Goal Conciliation Theory abductive-deductive architecture

Abduction	[1]		Q
	Deduction	[2]	P
		[3]	P
		[4]	Q'

Source: Rauen (2018, p. 14).

In this architecture, goal conciliation occurs whenever the consequent state Q' in stage [4] satisfies the expectations of the goal Q in stage [1] and goal non-conciliation conversely. Additionally, there is an active self-conciliation whenever the individual verifies whether the antecedent action achieves the consequent state Q' in stage [4], and active hetero-

6. Q' represents that the goal achievement tends to be slightly different from the projected goal Q .

conciliation—as in *Golden Balls*—whenever this verification involves more than one individual.⁷

Moreover, Rauen (2014) asserts that conciliations enhance the predictive power of an antifactual abductive hypothesis PQ , while non-conciliations diminish it. Depending on the level of prediction, a hypothesis can fall into five categories: categorical $P \Leftrightarrow Q$, indicating certainty; biconditional $P \leftrightarrow Q$, indicating necessity and sufficiency; conditional $P \rightarrow Q$, indicating sufficiency but not necessity; enabling $P \leftarrow Q$, indicating necessity but not sufficiency; and, finally, tautological $P \dashv Q$, indicating neither necessity nor sufficiency. Considering that Nick can only achieve Ibrahim's cooperation through communicative stimuli, which is not guaranteed, we can conclude that Nick employs an enabling antifactual abductive hypothesis.

Furthermore, since these utterances serve a practical goal or intention in Goal Conciliation Theory, analyzing communicative processes requires three layers of intention. We begin with at least one practical goal or intention Q , which governs an informative goal or intention P , which in turn governs a communicative goal or intention O , and this, in turn, governs the factual execution N of the plan—the utterance itself. Therefore, it is reasonable to assume that Nick's utterances (1–2) serve a superordinate intentional action plan guided by a practical goal or intention, with the most plausible hypothesis being that he initially intends to steal the £13,600 prize. Since Nick can only achieve this practical goal Q with Ibrahim's collaboration, he must inform Ibrahim of it (P) and communicate it (O) through an open ostensive stimulus (N).

We can see this intentional action plan in Figure 3:⁸

Figure 3 – Nick's intentional action plan

Q – Steal the £13,600 prize, Nick (practical goal)
P – Inform Ibrahim that Nick will certainly pick the steal ball in <i>Split or Steal</i> , Nick (informational goal)
O – Communicate to Ibrahim that Nick will certainly pick the steal ball in <i>Split or Steal</i> , Nick (communicational goal)
N – Utter “Ibrahim, I want you to trust me. One hundred percent I am going to pick the steal ball,” Nick (execution of the plan)

Source: Figure developed by the authors.

7. Given that the antecedent action P in stage [3] is optional, there is active (in)conciliation whenever the individuals perform it, and there is passive (in)conciliation conversely.

8. For practical purposes, we use descriptions with successively embedded goals.

Doubting the proposal, Ibrahim activates strategies of vigilance toward the content of the message.⁹ During the negotiation, he poses a set of questions, expecting Nick to confirm the explicature of his proposal.

Let us consider the first question:

(3) Ibrahim (02:44): Sorry, you gonna pick...?

Ibrahim wants to know if it is true that Nick is going to pick the steal ball with one hundred percent certainty¹⁰.

Rather than providing a simple ‘yes’ or ‘no’ answer to the question, Nick opts for a paraphrased version of the propositional content from his initial utterance. He replaces the verb ‘to pick’ with ‘to choose,’ expecting Ibrahim to draw the intended inference.

(4) Nick (02:45): I gonna choose the steal ball.

Nick says that Nick is going to choose the steal ball.

Nick suggests that it is true that Nick is going to pick the steal ball with one hundred percent certainty (implied conclusion by *modus ponens*).

Remaining skeptical, Ibrahim rephrases his check in (3), substituting Nick’s verb ‘to choose’ with ‘to take.’

(5) Ibrahim (02:45): You gonna take the steal.

Ibrahim wants to know if it is true that Nick is going to take the steal ball with one hundred percent certainty.

Nick refines his proposal:

(6) Nick (02:46): I want you to do split, and I promise you that I will split the money with you.

9. Do saying P imply believing P and does believing P imply P in this context? Thus, is it plausible that by saying that Nick is going to pick the steal ball, he sincerely believes and intends to claim the prize?

10. According to Sperber and Wilson (1995, p. 252), questions are interpretations of relevant true answers. Hearers of interrogative utterances would retrieve their logical form and integrate it into a description in the form “The speaker asks *WH-P*,” where *WH-P* is an indirect question. The authors distinguish “yes-no questions” and “*WH*-questions,” depending on whether they have a complete propositional form or do not.

Nick says that Nick wants Ibrahim to pick/choose/take the split ball.

Nick says that Nick promises that Nick will split the £13,600 prize with Ibrahim after the *Golden Balls* show.

The commitment to split the prize implies that Nick's intentional action plan is somewhat more intricate. Stealing the prize is not the primary practical goal, but serves the subsequent purpose of sharing the £13,600 prize after the show.

Figure 4 – Revised Nick's intentional action plan

Q – Share the £13,600 prize with Ibrahim after the show, Nick (higher-level practical goal)
P – Steal the £13,600 prize, Nick (lower-level practical goal)
O – Inform Ibrahim that Nick wants Ibrahim to pick the split ball and that Nick will split the £13,600 prize with him after the show (informational goal)
N – Communicate to Ibrahim that Nick wants Ibrahim to pick the split ball and that Nick will split the £13,600 prize with him after the show, Nick (informational goal)
M – Utter "I want you to do split, and I promise you that I will split the money with you" (execution of the plan)

Source: Figure developed by the authors.

3. Where is Nick's point coming from?

Stunned by Nick's proposal, Ibrahim develops a new series of checks to confirm it. Nick, in turn, provides the alleged evidence.

- (7) Ibrahim (02:50): After you took steal.
- (8) Nick (02:51): Yeah.
- (9) Ibrahim (02:52): You gonna take steal.
- (10) Nick (02:52): Yeah.
- (11) Ibrahim (02:53): I gonna take split.
- (12) Nick (02:53): Yeah.
- (13) Ibrahim (02:54): So, you take the money.
- (14) Nick (02:56): And I gonna split it with you.
- (15) Ibrahim (02:58): After the show.
- (16) Nick (02:58): Yeah.

The audience laughs and applauds. Nick adds:

- (17) Nick (03:07): Ibrahim I promise you I will do that.
- (18) Nick (03:09): If... If... If you do steal, we both walk away with nothing.¹¹
- (19) Nick (03:11): I'm telling you one hundred percent I'm gonna do it

11. Note how the utterance in (18) clarifies that Ibrahim will lose his £6,800 prize, choosing the steal ball. Explicitly, *Nick says that* if Ibrahim picks the steal ball, then Nick and Ibrahim will leave the *Golden Balls* show without anything of the £13,600 prize.

Ibrahim attempts to persuade Nick to reconsider: both could pick the split ball.

(20) Ibrahim (03:13): I appreciate that, but I'll give you another alternative.

(21) Ibrahim (03:18): Why don't each of us both pick split.

Nick maintains his proposal.

(22) Nick (03:20): I'm not gonna pick split.

(23) Nick (03:21): I'm gonna steal.

(24) Nick (03:22): Ibrahim, I honestly one hundred percent gonna steal.

Ibrahim checks whether it is true that it is in Nick's nature to steal, and Nick answers this yes/no question with "No".

(25) Ibrahim (03:24): It's in your nature to steal?¹²

(26) Nick (03:25): No.

According to Sperber et al. (2010), we can direct epistemic vigilance toward the source or the content of the information. Utterances from reliable informants must fulfill the conditions of competence and benevolence. The authors contend that competence involves the accurate transmission of information by the speaker. In contrast, benevolence refers to the speaker's intention to share truthful and authentic content, explicitly excluding disinformation, deception, or omission, regardless of whether these stem from malice or indifference. Other cases involve vigilance toward the risk of deception and falsehood.

When individuals direct their epistemic vigilance toward the risk of deception, this requires an understanding not only of the communicator's epistemic states but also of their intentions, including intentions to induce false beliefs in the audience. Such understanding demands a relatively sophisticated mind-reading through higher-order metarepresentations. For example, "*She believes that not-P but wants me to believe that P*" combines a first-order attribution of belief with a second-order attribution of intention (Sperber et al., 2010, p. 372).

12. Explicitly, *Ibrahim wants to know if it is true that it is Nick's nature to steal*.

According to Sperber et al. (2010), vigilance regarding the content of an utterance tends to foster the belief that some contents are inherently plausible, even if they originate from an unreliable source, while others are deemed unbelievable, even if they come from a reliable source. Nevertheless, they emphasize that epistemic vigilance directed at communicated content should—in most cases—extend beyond its inherent logical properties. People evaluate utterances by considering the credibility of the information within the context of their existing beliefs.

Nick addresses doubts about his honesty by affirming his integrity and promising to make Ibrahim aware of it. Ibrahim expresses skepticism, interrupting Nick's speech. In response, Nick reiterates his honesty and reveals that he intends to steal the prize precisely because of that honesty.

(27) Nick (03:27): I... I'm honest and I'm gonna tell you.

(28) Ibrahim (03:28): I think you are honest.

(29) Nick (03:28): I am.

(30) Nick (03:28): That's why I'm telling you I'm gonna steal.

(31) Nick (03:30): If you do split, then I will split the money.

Ibrahim cannot imagine picking the split ball while Nick chooses the steal one.

(32) Ibrahim (03:32): I can't... I can't see myself doing that.¹³

Supposedly because Ibrahim might be considering stealing, Nick adds:

(33) Nick (03:34): Okay.

(34) Nick (03:34): Well, I'm gonna steal.

Nick (03:35): So, we are going to leave with nothing.¹⁴

Ibrahim's proposal to dissuade Nick from his plan seems to have no effect. Thus, Ibrahim asks where Nick's brains come from [sic]. Nick points to his head, eliciting laughter from the audience.

(36) Ibrahim (03:37): Where's your brains coming from?

(37) Nick points to his head (03:38).

13. Explicitly, *Ibrahim says that* Ibrahim cannot see Ibrahim picking the split ball while Nick picks the steal ball to share the £13,600 prize with Ibrahim after the Golden Balls show.

14. Explicitly, *Nick says that* if Ibrahim and Nick pick the steal ball, then Ibrahim and Nick will leave the Golden Balls show with nothing of the £13,600 prize.

Next, Ibrahim expresses that he still does not trust Nick's proposal. In response, Nick reaffirms his honesty and reiterates his commitment to splitting the prize with Ibrahim after the show. Despite Ibrahim's insistence that they both select the split ball, Nick remains firm, declaring that he will choose the steal ball.

- (38) Ibrahim (03:42): I can't work that.
- (39) Nick (03:44): I know that I'm a decent guy, and I will split the money with you.
- (40) Ibrahim (03:48): I thought we both should split them.
- (41) Nick (03:48): No, I'm gonna do steal.

Surveying the conversation, Jasper Carrott warns Ibrahim that there is no legal requirement for Nick to be bound to fulfill his promise, and Ibrahim says that he knows that, overlapping with Carrott's words.

- (42) Jasper Carrott (03:51): There is no legal requirement for him to give you the money.
- (43) Ibrahim (03:51): No, I know that, of course.

Aware that the *Golden Balls* program does not guarantee Nick's proposal will be honored, Ibrahim urges Nick to keep his word if he makes a promise, recalling lessons on honor passed down from his father.

- (44) Ibrahim (03:56): If I might give my word...
- (45) Ibrahim (03:57): Now let me give you a glimpse of what my word means.
- (46) Nick (03:59): OK.
- (47) Ibrahim (04:00): My father once said to me.
- (48) Ibrahim (04:02): A man who does not keep his word is no man.
- (49) Ibrahim (04:06): He's not worth nothing.
- (50) Ibrahim (04:07): [...] not worth a dollar.
- (51) Nick (04:08): I agree.
- (52) Ibrahim (04:10): So...

Nick interrupts Ibrahim's argument, reiterating that he will steal the prize and emphasizing that the choice is Ibrahim's. He warns that if Ibrahim also chooses the steal ball, both will lose the £13,600 prize.

- (53) Nick (04:11): Ibrahim, I'm gonna steal, so you've got the choice.
- (54) Nick (04:14): You either steal, and we both walk away with nothing, because, you know, I've told you my intention, and I've told you that I will split the money with you, Ibrahim.

Nick assures Ibrahim that he will keep his word. Consequently, he places the responsibility on Ibrahim to choose between the split ball—ensuring the prize will be divided, provided Ibrahim receives his share after the game—and the steal ball, which would result in both players leaving with nothing. After all, Nick has already disclosed his intention. Ibrahim's response suggests that he is inclined to pick the split ball, contingent on Nick honoring his promise to share the money after the show.

(55) Ibrahim (04:20): If I gave my word that I was gonna split.

(56) Ibrahim (04:21): I would split and you gonna take steal, so...

Jasper Carrott intervenes again, urging the players to make their own decisions. He emphasizes that, according to the *Golden Balls* rules, the only way to guarantee half the prize is for both to choose the split ball—a point with which Ibrahim concurs.

(57) Jasper Carrott (04:26): The only way you can guarantee to walk away with £6.800... it's the guarantee that you both put the split ball in.

(58) Ibrahim (04:31): It's both sharing it.

(59) Jasper Carrott (04:35): And I do now have to push you for a decision.

(60) Jasper Carrott (04:37): It's a tough one.

At this moment, Ibrahim shows signs that he is going to steal the prize.

(61) Ibrahim (04:38): We've lost it.

(62) Ibrahim (04:38): We've lost everything.

(63) Nick (04:38): I think we lost it.

As Nick confirms that both are losing the prize, it becomes apparent that Ibrahim is in a deadlock. Recognizing that his attempts to change Nick's standpoint have been ineffective, Ibrahim accuses his opponent of being an idiot for proposing to take the money he believes is his. When Ibrahim says Nick is walking away with his money, he also suggests that he could pick the split ball, despite considering that Nick may not fulfill the promise to share the stolen amount after the game.

(64) Ibrahim (04:41): You are walking away with my money because you are an idiot.¹⁵

15. We can describe both inferences as follows:

A₁ – Ibrahim says that Nick is walking away with Ibrahim's money because Nick is an idiot [for proposing to Ibrahim to split the £13,600 jackpot while Nick steals the £13,600 jackpot

Nick disagrees, but Ibrahim continues to call him an idiot.

- (65) Nick (04:43): No, that's not right.
- (66) Ibrahim (04:44): You're an idiot.
- (67) Ibrahim (04:44): You're an idiot.
- (68) Ibrahim (04:45): That was right, you're an idiot.
- (69) Ibrahim (04:47): That's what you are.

Jasper Carrott intervenes once again, urging both participants to make their choices.

- (70) Jasper Carrott (04:49): Wait, this can go on all night, and these people gonna get up for breakfast.
- (71) Jasper Carrott (04:54): Nick, choose to split or steal
- (72) Jasper Carrott (04:57): Ibrahim, choose split or steal.
- (73) Jasper Carrott (05:00): Now, please, choose the ball.

Both choose the ball on the right, but Ibrahim suddenly swaps his ball, saying that he is going with Nick and suggesting that he will pick the split ball.

- (74) Ibrahim (05:03): Well, I'll tell you what... I'm gonna go with you.
- (75) Nick (05:05): Okay.
- (76) Ibrahim (05:05): I'm gonna go with you.¹⁶

S_1 – *Ibrahim says that* Ibrahim is telling Nick that Ibrahim is going with Nick's proposal.

S_2 – Ibrahim *probably* will pick the split ball (implied conclusion by *modus ponens* $S_1 \rightarrow S_2$).

Nick, once again, restates his promise.

- (77) Nick (05:07): I've promised you I'm splitting.

and promises to share the £13,600 jackpot with Ibrahim after the Golden Balls show] (implied premise from the explication of (64)).

A_2 – Ibrahim probably considers that Nick may not fulfill his promise to share the £13,600 jackpot with Ibrahim after the Golden Balls Show (implied conclusion by *modus ponens* $A_1 \rightarrow A_2$).

A_3 – Ibrahim could probably pick the split ball (implied conclusion by *modus ponens* $A_1 \rightarrow A_3$)

16. We can describe the inference as follows:

A_1 – *Ibrahim says that* Ibrahim is telling Nick that Ibrahim is going with Nick's proposal (implied premise from the explication of (76)).

A_2 – Ibrahim probably will pick the split ball (implied conclusion by *modus ponens* $A_1 \rightarrow A_2$).

Jasper Carrott requests that they reveal their chosen balls.

(78) Jasper Carrott (05:12): You cannot change your balls now.

(79) Jasper Carrott (05:17): Split or Steal.

Both choose the split ball (05:21).

Figure 5 – Ibrahim and Nick show the split ball



Source: Screenshot from YouTube video *The weirdest split or steal ever!* (Spinout3, 2010).

4. What happened here?

At this point, it is worth asking whether Nick regretted his strategy or whether the strategy was different. Note that Nick's utterance in (77) is highly ambiguous.¹⁷ Until now, we assumed that the explicature of this utterance had to do with the promise of sharing the prize afterward. However, Nick's choice of the split ball suggests another interpretation—that the highest-level practical goal could be to split the prize.

If Nick wanted to divide the jackpot from the beginning, why did he say he would steal the prize? We argue that Nick would dissuade Ibrahim from the temptation to steal with this strategy. Thus, what seems to have happened is that Nick's proposal *O* was in the service of a sub-goal *P* of

17. We could describe this ambiguity as follows:

A_1 – Nick says that Nick promised that Nick would split the £13,600 prize with Ibrahim after the show.

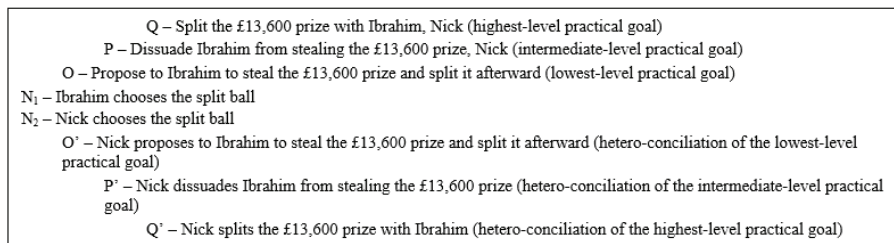
A_2 – Nick says that Nick promised that Nick would split the £13,600 prize with Ibrahim during the show.

dissuasion, and this sub-goal *P* of dissuasion was in the service of the highest goal *Q* of splitting the prize.

In this context, for Nick's plan to yield the desired effects, there was no room for hesitation, as any wavering from Nick could encourage Ibrahim to choose the steal ball.

Let us consider:

Figure 6 – Achievement of Nick's revised intentional action plan



Source: Figure developed by the authors.

In sum, Nick supposedly wanted the division of the prize, the best alternative for both participants in Golden Balls (or the *Nash Equilibrium* in the context of the *Prisoner's Dilemma*). With Nick's proposal, Ibrahim spends the entire negotiation in a false dilemma between picking the split ball, trusting Nick's promise, and picking the steal ball, risking that both would lose the prize.

Nick employs an audacious and intriguing strategy. Ibrahim never considers the hypothesis that Nick was bluffing because picking the split ball seems to be the only way Ibrahim could take his £6,800 prize. Ibrahim, Jasper Carrott, the audience, and, arguably, first-time readers or YouTube watchers cannot see the countermove, supposedly because of the possibility of losing everything. All Ibrahim had to do was choose the steal ball, and he would have walked away with the entire £13,600 jackpot.

Let us consider how we could model this countermove:

Figure 7 – Potential non-conciliation of Nick's revised intentional action plan

Q – Split the £13,600 prize with Ibrahim, Nick (highest-level practical goal)
P – Dissuade Ibrahim from stealing the £13,600 prize, Nick (intermediate-level practical goal)
O – Propose to Ibrahim to steal the £13,600 prize and split it after the show (lowest-level practical goal)
N ₁ – Ibrahim chooses the steal ball
N ₂ – Nick chooses the split ball
O' – Nick proposes to Ibrahim to steal the £13,600 prize and split it after the show (hetero-conciliation of the lowest-level practical goal)
–P' – Nick does not dissuade Ibrahim from stealing the £13,600 prize, Nick (hetero-non-conciliation of the intermediate-level practical goal)
–Q' – Nick does not split the £13,600 prize with Ibrahim, Nick (hetero-in conciliation of the highest-level practical goal)

Source: Figure developed by the authors.

5. Final remarks

We explore in this article how Nick and Ibrahim negotiate the £13,600 jackpot in the *Split or Steal* round of the British game show *Golden Balls*, synergistically adopting the descriptive-explanatory resources of Relevance and Goal Conciliation Theories. The results support the hypothesis of an asymmetric relationship in this negotiation. Nick proposes an intentional action plan by designing, executing, and monitoring antefactual abductive hypotheses that enable him to collaboratively manage the achievement of his practical goals (Rauen, 2014). Ibrahim mobilizes practical-epistemic vigilance strategies (Sperber et al., 2010; Luciano, 2019) to infer from the ostensive stimuli (Sperber & Wilson, 1995) what these goals were and, thus, to make his decision.

As we demonstrated, Goal Conciliation Theory allowed us to analyze the proposer's intentional action plan, that is, how Nick placed Ibrahim in the contingency of a dilemma between choosing the split ball and trusting Nick's word that he would share the prize with Ibrahim after the show, or both choose the steal ball and leave with nothing. Furthermore, we could observe how Relevance Theory allowed us to describe and explain how Ibrahim (and the audience) fitted Nick's propositions into logical forms with which they paired linguistic and encyclopedic inputs, generating explicatures and subsequent implicatures that constitute the proper dilemma.

In the episode, Nick takes the lead and proposes to Ibrahim that he will steal the prize with one hundred percent certainty. Confronted by such an unusual proposal, Ibrahim produces a series of checks, accuses his opponent of being an idiot, and appeals to arguments of honor, but finally accepts the proposal, convinced of Nick's sincerity. To everyone's surprise, Nick also

chooses the split ball, suggesting that his intentional action plan included an intriguing strategy to ensure the prize was split by preventing Ibrahim from stealing.

Ibrahim's inability to recognize Nick's strategy indicates that the stressful conditions of negotiation in *Split or Steal* may compromise practical-epistemic vigilance and prevent the formulation of appropriate countermeasures. Therefore, it is not surprising that readers are led to believe—given Nick's assertiveness—that he will share the prize after the show. We only realize after the game that the proposal was a dissuasion strategy that Nick had implemented despite the risk that Ibrahim could calculate the bluff and steal the prize.

The episode featuring Nick and Ibrahim—an illustration of optimal strategic success or a *Nash Equilibrium* in Game Theory, where both contestants walk away with half of the prize—sheds light on how players construct and monitor intentional action plans. Epistemic vigilance about the truthfulness of information and the reliability of its sources is not sufficient. One must also remain alert to how individuals use this information, particularly when self-interest, rather than altruism, is involved. In this case, both participants were sincere. However, as Belan (2022) observes, this is not always the case, highlighting the need for further investigation into the effectiveness of such monitoring. In an age marked by the deliberate spread of disinformation, this remains a pressing and relevant issue.

Acknowledgments

The authors express their gratitude to the Graduate Program in Language Sciences of the University of Southern Santa Catarina for its institutional support. Fábio José Rauen thanks the Ânima Sociesc Institute for Innovation, Research, and Culture for awarding the Research Productivity Grant. Andressa Bregalda Belan thanks the Federal Institute of Santa Catarina (IFSC) for granting a full leave, as well as the Coordination for the Improvement of Higher Education Personnel (CAPES)—Financing Code 001—and the UNIEDU/FUMDES Program for awarding doctoral scholarships.

Conflict of interests

The authors declare that they have no conflict of interests.

Credit Author Statement

We, Fábio José Rauen, and Andressa Bregalda Belan, have all participated in study conceptualization, methodology, analysis, and writing. All authors approve the final version of the manuscript and are responsible for all aspects, including the guarantee of its veracity and integrity.

Data Availability

The data used in this study is available in the YouTube video “Golden Balls. The Weirdest Split or Steal Ever!” and can be accessed at <https://www.youtube.com/watch?v=S0qjK3TWZE8&t=51s>.

References

- Belan, A. B. (2022). *Golden Balls: Condições de negociação e vigilância prático-epistêmica* [Doctoral dissertation, University of Southern Santa Catarina]. <https://repositorio.animaeducacao.com.br/handle/ANIMA/25761>
- Bêrni, D. de A., & Fernandez, B. P. M. (2014). *Teoria dos jogos: Crenças, desejos e escolhas* [Game theory: Beliefs, desires, and choices]. Saraiva.
- Fiani, R. (2015). *Teoria dos jogos* [Game theory]. Elsevier.
- Luciano, S. F. M. (2019). *Vigilância epistêmica e prática: Uma abordagem orientada pelo conceito de conciliação de metas* [Epistemic and practical vigilance: An approach guided by the concept of goal conciliation] [Doctoral dissertation, University of Southern Santa Catarina]. <https://repositorio-api.animaeducacao.com.br/server/api/core/bitstreams/876f66fa-ff5d-442b-87f6-ce04a3a20bf6/content>.
- Rauen, F. J. (2008). Processos interacionais discente/docente em espaço virtual de aprendizagem: Análise com base na teoria da relevância [Student/teacher interaction processes in a virtual learning environment: Analysis based on relevance theory]. *Scripta*, 22, 190-217. <https://periodicos.pucminas.br/scripta/article/view/4440/4597>
- Rauen, F. J. (2014). For a goal conciliation theory: Ante-factual abductive hypotheses and proactive modelling. *Linguagem em (Dis)curso*, 14(3), 595-615. <https://doi.org/10.1590/1982-4017-140309-0914>
- Sierralta Ríos, A. (2017). *Negociação e teoria dos jogos* [Negotiation and Game Theory]. Thomson Reuters.
- Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd ed.). Blackwell. (Original work published 1986)
- Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., and Wilson, D. (2010). Epistemic vigilance. *Mind & Language*, 25(4), 359-393. <https://doi.org/10.1111/j.1468-0017.2010.01394.x>

- Spinout3. (2010). *The weirdest split or steal ever!* [Video]. YouTube. <https://www.youtube.com/watch?v=S0qjK3TWZE8&t=51s>.
- Wegmüller, M. (2020). A man, his words and his two golden balls: Game theory in practice. *HEconomist*, 9(1), 16-19. <https://heconomist.ch/2020/10/26/a-man-his-words-and-his-two-golden-balls-game-theory-in-practice/>.
- Wikipedia. (n.d.). *Golden Balls*. Retrieved January 5, 2024, from https://en.wikipedia.org/wiki/Golden_Balls.
- Wilson, D. (2004). *Pragmatic theory*. UCL.

Recebido em: 21.02.2024

Aprovado em: 08.07.2025

EDITORA GERAL: Marcia Veirano Pinto 