THE EFFECTS OF INFORMATION PRESENTATION AND SITUATIONAL COMPLEXITY ON EFFECTIVE DECISIONMAKING

Stephen D. Donnel, Department of Operational Sciences, Air Force Institute of Technology, 2950 Hobson Way, Wright-Patterson AFB, OH 45433, 937-255-6565, stephen.donnel@us.af.mil

William N. Caballero, Department of Mathematical Sciences, United States Air Force Academy, 2354 Fairchild Drive, USAF Academy, CO 80840, USA, william.caballero@afacademy.af.edu

Brian J. Lunday, Department of Operational Sciences, Air Force Institute of Technology, 2950 Hobson Way, Wright-Patterson AFB, OH 45433, 937-255-6565 x4624, brian.lunday@afit.edu

ABSTRACT

This research examines and quantifies the effects of both information presentation modality and situational complexity on an individual's ability to determine an effective strategy in a competitive environment. An empirical study presents a series of two-player, normal-form games with complete information to a set of participants and asks participants to identify their best response for each game. The games vary both in their complexity and their manner of presentation to participants. Dual channel theory directly informs the experimental design; it specifies how humans process information and posits the existence of capacity limits to such cognitive mechanisms. Aside from their relevance to descriptive decision theory, the insights derived from this analysis provide utility to a wide array of decisionmaker-centered tools and techniques.

Keywords: Descriptive Decision Theory, Multimedia Learning, Behavioral Game Theory, Cognitive Hierarchy, Multiple Correspondence Analysis

INTRODUCTION

The motivation behind this research is to explore and statistically characterize how information presentation and situational complexity affect an individual's ability to reach sound, strategic decisions in a competitive environment. This outcome is accomplished via a designed experiment that asks members of a test population to make utility-maximizing decisions in normal-form games under varied conditions. In so doing, this study seeks to characterize the relative quality of strategic decisions over different aspects of the test.

The challenge of effectively communicating complex information is well-studied in educational psychology and, although scholars in this field tend to focus on information retention and cognitive processing, their qualitative insights directly inform our study. Namely, the Cognitive Theory of Multimedia Learning (Mayer 2009) describes how information presentation affects *learning* and provides principles that educators can utilize to maximize their effectiveness. We examine herein the extent to which these principles extend to *strategic decision-making* by including them as factors within a designed experiment. Herein, *strategic* refers to an environment in which multiple decision-makers will seek to maximize their respective utilities, which depend upon the collective decisions made. This context differs from alternative frameworks (e.g., see Terzopoulou & Endriss 2022) wherein an individual attempts to manipulate collective decisions in their own self-interest.

To explore the effect of these factors on strategic decision-making, one must first ascertain a means to characterize decision quality. Such an endeavor is non-trivial; however, for specific classes of competitive interactions, game theory provides well-established concepts to describe strategic reasoning. As such, within our experimentation, participants play a set of normal-form games wherein a unique best response exists. Therein, the best response is identifiable by either an iterative best response analysis or the iterative elimination of dominated strategies. In short, there exists for each game a "correct" action which the participants must try to identify. The existence of a correct answer is unlike most studies in the field of neuroeconomics, wherein much of the focus is on either perceptual or value-based decisions (Serra 2021). Nevertheless, such a binary metric of "correct" or "incorrect" cannot fully address the nuanced nature of strategic decision-making. Even though alternative decisions may not select the ideal action, a relative preference ordering among actions may be established, which can be used to infer varying levels of strategic competency. Behavioral game theory provides a conceptual framework to quantify such subtleties. In particular, the cognitive hierarchy model set forth by Camerer et al. (2004) can characterize the performance of a population on a normal-form game based upon their average level of strategic thought. Higher levels of strategic thought imply more sophisticated decision-making. Therefore, this study uses as response variables both best-response identification and level of strategic thought.

Such disparate response variables used to quantify decision quality require distinct means of statistical analyses. Herein, analysis initially conducts multiple forms of categorical data analyses without any game theoretic assumptions to explore associations between information presentation, situational complexity, and decision quality. Application of the Generalized Cochran-Mantel-Haenszel (GCMH) test and multiple correspondence analysis (MCA) collectively examine the extent to which the theory of multimedia learning extends to strategic decision-making. The GCMH test ascertains whether differing forms of information presentation affect the respondent's decisions, i.e., whether their respective marginals across actions are homogeneous (Agresti 2002). MCA identifies more nuanced associations in the data, i.e., to determine the degree that participant behavior across information-presentation and complexity levels is coincident or divergent (Greenacre & Blasius 2006). Additional analysis under behavioral game theoretic assumptions provides further insight into the players' behavior. This analysis requires a grid search for point estimation of the participants' average level of strategic thought and resampling methods to characterize uncertainty about this point estimate (Efron & Tibshirani 1994).

In a competitive environment, information presentation should promote understanding to maximize the efficacy of strategic thought. Therefore, as the experimentation and analysis provided herein seeks to identify trends and insights regarding the impact of information presentation and situational complexity on strategic behavior, the resulting insights also inform the practice of such operations research in competitive environments.

BACKGROUND

Dual Coding Theory and Multimedia Learning The educational, psychological concept of *generative cognitive processing*, i.e., an individual's ability to understand presented information (Mayer 2010), is of central importance to this research. More specifically, the information and cognitive processing mechanisms proposed by both the dual coding and multimedia learning theories serve as conceptual underpinnings for the communication principles tested herein.

To determine how presentation modality may affect strategic decision-making, it is necessary to first consider how humans encode information psychologically. Dual coding theory asserts that humans use both visual and verbal information to represent knowledge. Visual and verbal information are each respectively received and processed within two cognitive subsystems. These subsystems are used to process incoming information from non-verbal occurrences and language. That is, analog codes store images, and symbolic codes store letters and words. The validity of these assertions was tested by Paivio (1969), who studied an individual's ability to recall images and words in sequential order. Within Paivio's study, it was observed that participants could recall the sequential order of words more readily than the ordering of images. Such results support dual coding theory's hypothesis that verbal and visual information are processed distinctly.

The Cognitive Theory of Multimedia Learning set forth by Mayer (2009) suggests that one may need to systematically balance information presentation to enable decision maker success.. Akin to dual coding theory, it assumes that audio and visual information are processed separately, abstracted distinctly in working memory, and subsequently integrated with prior knowledge for information comprehension. By additionally assuming that the dual channels have limited capacity and that humans must engage with material to learn, Mayer sets forth a set of information-presentation principles that promote knowledge acquisition. The modality, multimedia, and personalization principles are conspicuously relevant. The modality principle asserts that comprehension is promoted when visual cues are paired with spoken words. The multimedia principle stresses that comprehension is promoted when pictures are paired with printed words instead of relying upon printed words alone. Furthermore, the personalization principle indicates that conversational language engenders effective cognitive processing better than formal prose (Mayer 2009).

Should these principles indeed promote knowledge acquisition, they provide a natural foundation upon which to base information presentation. However, the degree to which the Cognitive Theory of Multimedia Learning extends beyond cognitive processing and into strategic decision-making is understudied. To wit, there are studies that examine dual coding theory principles vis-à-vis effective information *processing* and *recall* (e.g., see Vaid 1988, Clark & Paivio 1991, Mayer 2009, DeLong et al. 2021), but we have not found any foundational work to evaluate the relative efficacy of information presentation when it must be *applied* to inform strategic decisions. We contend that the effectiveness of communication must consider the quality of the decisions, not simply the ability of a decision maker to remember the information. Therefore, the modality and multimedia principles are leveraged to construct our test instrument and experimental design. Their interaction with additional factors (i.e., situational complexity) is also explored. The initial hypothesis suggests that the most pronounced effect of these principles presents under conditions of greater complexity and lesser familiarity; however, the formal examination of this hypothesis is a primary contribution of this manuscript.

Assessing Strategic Competence in Normal-Form Games Given each player has complete information on A and u, normative solution concepts consider Player i's *best response* to their opponents' strategy profile s_{-i}. A strategy s_i for Player i is a best response to s_{-i} if no alternative strategy yields a higher expected utility for Player i. A Nash equilibrium is a set of mutual best responses and, should a normal-form game have only one such equilibrium, it is said to *solve* the game. A player's ability to identify such an equilibrium therefore reflects positively on their ability to reason strategically. Unfortunately, multiple equilibria may exist, in general, thereby complicating Nash-equilibrium identification as a measure of strategic thought.

Alternatively, if dominance relationships exist among a player's strategies, analysis of a perfectly rational player may be greatly simplified. A strategy s_i is said to dominate another strategy s'_i when s_i always yields utility greater than or equal to s'_i for any s_{-i}, and it strictly dominates s'_i when the inequalities are strict. A strategy for Player i is *strictly dominant* when it strictly dominates all other strategies. Conversely, a strategy s_i is *strictly dominated* if another s'_i always yields greater utility. From a normative perspective, strictly dominated strategies can be removed from consideration because a rational player will never select them over a non-dominated strategy. Therefore, when dominance relationships exist, the selection of a dominated strategy (or non-selection of a dominating strategy) can be used to infer the quality of a strategic decision. Moreover, if each player has a dominant strategy, the associated joint strategy is necessarily a Nash equilibrium(Shoham & Leyton-Brown 2008).

Although dominance is a useful tool for assessing decision quality, analysis relying upon it is dichotomous. A dominant action is either selected or it is not. Nevertheless, based upon the extensive experimentation of behavioral game theorists, humans are known to exhibit some degree of bounded rationality (Camerer 2011), often departing from a Nash equilibrium strategy in practice (Alsaleh & Sayed 2022). One should expect to observe some "irrational" strategies, but not all such behavior is equivalent. Studies have investigated team influence on decision-making (Arad et al. 2022) and the personal belief of trust in others to make sound, reasonable decisions (Li et al. 2019). Some reasoning errors may be associated with a higher degree of strategic sophistication than others. Status and rank in the corporate hierarchy exhibit different levels of strategic decision-making to increase earning potential (Holm et al. 2020). To more completely measure a sampled population's degree of strategic thought, we turn to behavioral game theory for additional analysis.

In particular, this research relies upon the cognitive hierarchy (CH) model, an extension of the traditional Level-k (LK) model, to evaluate such nuance. Both solution concepts describe empirical game play by assuming players are defined by a "level" or "step" of strategic thought that determines the depth of strategic reasoning they can perform. For example, in the LK model, a level-0 player is assumed to choose an action randomly. In turn, believing that all their opponents are level-0 thinkers, a level-1 player selects a response that maximizes their expected utility. Level-2 players act similarly, but believe all their opponents are level-1 thinkers; level-3 players select an expected-utility maximizing action believing that all their opponents are level-2 players, and so on. Thus, a level-k player always believes themself to be the most strategic player in the game and their opponents to be level-$(k-1)$ thinkers (Stahl & Wilson 1995).

The CH model extends the LK model by allowing step-k players to hold more realistic beliefs about their competitors (Camerer et al. 2004). Of note, Cognitive Hierarchy uses the term *step* instead of *level* to distinguish its characterizations linguistically from terms utilized in the Level-k modeling framework. Whereas, step-0 players still make no assumption about their opponents and select an action based on a uniform probability distribution over the action space, a step-k thinker (i.e., which differs from a level-k thinker) does not assume all others are step-$(k-1)$ thinkers. Rather, a step-k thinker (with $k > 0$) believes their opponents' levels of strategic thought range from step-0 to step-$(k-1)$, where the distribution over those levels follows a normalized Poisson distribution. The Poisson distribution exhibits parsimony via a single parameter, τ, i.e., the mean level of thought. As τ increases, the normalized frequencies from the step-k thinker's perspective impose overwhelming weight on $k-1$, so decisions within the CH modeling construct converge to a Nash equilibrium as $\tau \to \infty$ in games of iterated dominance.

Since τ represents the average level of strategic thought exhibited by a population in a competitive interaction, its estimation provides a natural measure of decision-making quality. Previous research in distinct applications may be used to gauge relative performance. For example, high τ-values were estimated in selected experiments to evaluate the strategic thought of professional game theorists and stock market portfolio managers (Camerer et al. 2004). Alternatively, across diverse subject pools and in different experimental games, nearly half of the 90 percent confidence intervals include $\tau = 1.5$, leading (Camerer et al. 2004) to suggest this value as a reasonable assumption for τ in untested games. These values provide benchmarks that can be used to assess the relative effect of information presentation on strategic decision-making quality under varied levels of situational complexity.

EXPERIMENT DESIGN AND ANALYTIC METHODOLOGY

Test Instrument Design This research explores the effect of information presentation and situational complexity on strategic decision-making. To analyze related hypotheses, this study designed, fielded, and analyzed the results of a computer-based test instrument. Namely, the study fielded an 18-question examination designed using Microsoft's Forms software, requiring participants to partake in a series of two-player, normal-form games. The test instrument presented games using varying combinations of presentation modalities to test the extension of the modality and multimedia principles to inform effective decision-making in situations of differing complexity. The set of identical questions for each participant manifested combinations of three presentation-modalities and four situational-complexity levels. Whereas one may describe situational complexity in many ways, we use game size as a measure because it directly corresponds to the relative amount of information one must synthesize to identify a high-quality decision.

Table 1 presents the test instrument structure, including, for each question, the size of the game, the presentation modality, and relevant notes. The examination begins with the Administration block to ensure participant comprehension and to comply with human-subject testing protocols. Question 1 presents the informed consent form per compliance with the Air Force Institute of Technology's Institutional Review Board, and Question 2-4 prepare the participant to complete the examination. Question 2 informs the participant to refrain from revisiting previous questions and changing their answer to prevent later experience from interfering with earlier responses. Question 3 is an audio check ensuring equipment calibration. Question 4 presents an example 2×2 game (i.e., $|A_i| = 2$ for both players) that introduces the language and presentation modalities used throughout the experiment.

Table 1: Test Instrument Design

Block	**Question**	**Scenario Size**	**Presentation Modality**	**Notes**
Admin.	1	-	-	Informed Consent
	2	-	-	Examination Instructions
	3	-	-	Audio Check
	4	-	-	Sample Game Orientation
Block 1	5	2×2	Audio	Iterative Best Response†
	6	3×3		IEDS‡
	7	4×4		Iterative Best Response†
	8	5×5		IEDS‡
Block 2	9	2×2	Visual	Iterative Best Response†
	10	3×3		IEDS‡
	11	4×4		Iterative Best Response†
	12	5×5		IEDS‡
Block 3	13	2×2	Audio & Visual	Iterative Best Response†
	14	3×3		IEDS‡
	15	4×4		Iterative Best Response†
	16	5×5		IEDS‡
Misc.	17	-	Audio & Visual	p-beauty contest style game
	18	-	-	Confirm directions followed

† Solvable by iteratively considering each player's best response(s)
‡ Solvable by *only* iteratively eliminating strictly dominated strategies

Blocks 1-3 use three different presentation modalities (i.e., audio-only, visual-only, and audio-and-visual) to present strategic situations of four different sizes (i.e., 2×2, 3×3, 4×4, and 5×5 normal-form games). These factor levels and the combinations thereof were selected to determine how multimedia-learning principles extend to effective strategic decision-making in environments of varied complexity. For each of the associated questions, the participant is asked to select an action that maximizes their utility, given that their opponent is self-interested. In the 2×2 and 4×4 scenarios, the best response is ascertainable by iteratively considering the players' best response choices. The 3×3 and 5×5 scenarios are rationalizable by eliminating strictly dominated strategies. Within our experiments, a game's utilities and its dominant action were varied across situational-complexity and presentation-modality levels. However, ordinal utility relationships between joint actions were maintained to ensure equitable comparisons of similarly sized games (i.e., the games' payoff structures are similar). For increased intelligibility, we henceforth label the dominant actions as number 1 in every game. Likewise, for analysis purposes, we number the remaining actions consistently according to their ordinal utility structure.

The examination terminates with the Miscellaneous block comprised of a benchmark question and an integrity check. Question 17 asks participants to play a commonly studied problem in the literature, the p-beauty contest game, to baseline their performance against other studies. This game requires participants to select a number between zero and 100 in an attempt for one's selection to be the closest to two-thirds of the average response; an audio-visual presentation is provided. Question 18 serves as an integrity check on the instructions presented in Question 2. The participant answers whether they returned to any previous question and adjusted their answer. Should this occur, the participant's answers are excluded from analysis.

Volunteers for the experiment were solicited from students and faculty in the Graduate School of Engineering and Management at the Air Force Institute of Technology. In total, 76 participants completed the test over approximately one month of data collection. Among the respondents, all students were active duty US Air Force (USAF) officers, and the faculty were a mix of civilians and active duty USAF officers.

Statistical Analysis Methodology for Raw Response Data An empirical data set can provide myriad insights, depending upon underlying assumptions about what generates the observed data. There exist multiple structural models in normative and behavioral game theory that have merit, but an assumption-free analysis of results should occur first. In this section, we describe a set of categorical data analysis techniques utilized on our raw response data to determine whether associations exist across factors, absent any additional game-theoretic assumptions.

The Generalized Cochran-Mantel-Haenszel (GCMH) test evaluates the homogeneity of the marginal distributions over the actions across each presentation modality. If marginal homogeneity is refuted, it can be inferred that differing presentation modalities coincide with behavior that is statistically distinct for games having the same complexity. Although the GCMH test is designed for discerning conditional independence, Agresti (2002) illustrates how it can be leveraged to answer questions regarding marginal homogeneity. If conditional independence is refuted, so is marginal homogeneity.

Multiple correspondence analysis (MCA) is a technique used to detect underlying structures in a data set comprised of nominal categorical data (Greenacre & Blasius 2006); it is computationally equivalent to correspondence analysis (CA) but allows one to analyze a pattern over several dependent variables. An MCA is constructed from I observations of K nominal variables wherein each variable has J_k levels and $J = \sum_{k=1}^{K} J_k$. Herein, a separate MCA is performed for each game size such that an observation is associated with each participant, the nominal variables are the presentation modalities, and the levels are the corresponding actions. The observed data is then summarized in a $I \times J$ indicator matrix, denoted by $\mathbf{X}$, that captures all the participants' data. MCA is subsequently performed by decomposing $\mathbf{X}$ or $\mathbf{B} = \mathbf{X}^T\mathbf{X}$ (i.e., the Burt Matrix) into a low-dimensional Euclidean space to identify associations.

Relationship identification with MCA is similar to CA and is often based on point proximity. Interpretation techniques may vary depending upon the coordinate system utilized; however, proximities are generally only meaningful between points in the same set (i.e., comparing among row or columns factors, respectively). As discussed by Greenacre & Blasius (2006), asymmetric maps can be used, to some degree, for comparing row and column points. Qualitative analysis can be utilized to identify the factor associated with each dimension, but the quality of these conclusions is limited by the MCA's percentage of total inertia, i.e., the square of the data's ϕ-coefficient explained by the selected dimensions (Greenacre & Blasius 2006). The ϕ-coefficient is a measure of association in the data where $\phi = \sqrt{\frac{\chi^2}{N}}$, χ^2 is the Pearson χ^2-statistic, and N is the total number of observations. Moreover, when performing MCA with either $\mathbf{X}$ or $\mathbf{B}$, it is well-known that the total inertia is artificially inflated via the inclusion of artificial variables and self-comparisons, respectively. As such, the calculated percentage of total inertia in most software packages is a lower bound on its true value. For more detail on CA and MCA, we refer the reader to research by Greenacre & Blasius (2006).

TESTING, RESULTS, AND ANALYSIS

Over four weeks, 76 participants took part in the experiment, completing the examination in approximately 47 minutes, on average. Three participants admitted on Question 18 that they violated the examination's rules, and their responses were removed from analysis.

The participants selected the proper action more often than not. However, responses implies varied behavior across presentation modalities. Given the ordinal payoff structure is constant across similarly sized scenarios, such a result suggests presentation modality may indeed affect strategic reasoning. Moreover, varied behavior across scenario sizes for similar presentation modalities indicates situational complexity is a complicating factor as well.

Prior to examining the treatment effects on the player's strategic reasoning, it is first worthwhile to benchmark the participants against others examined in the literature. Given its widespread use in the behavioral game theory literature, the p-beauty contest game is an ideal means to do so. For our sample population, the average response for the p-beauty contest was 31.02. This value corresponds to $\tau \approx 1.65$. Camerer et al. (2004) observed a median estimate for τ of 1.61 across 24 related, behavioral studies. Therefore, in terms of their strategic-reasoning ability, our population exhibits similar aggregate behavior to others examined in the literature.

Raw Response Analysis: Testing for Marginal Homogeneity Analysis reveals that presentation modality induced significantly different behavior across games of varied complexity. Table 2 displays the calculated GCMH test statistics for each scenario size. In each case, the associated p-value is less than 0.01, indicating a significant difference existed in how the participants responded between at least two of the presentation modality treatments for each scenario size. These results confirm the relevance of the dual-coding-theory construct to strategic environments.

Table 2: GCMH Test Statistics

2×2	3×3	4×4	5×5
45.3	190	167	111

With a significant difference identified across each scenario, the next phase of analysis conducted pairwise statistical tests on the presentation modalities for the different combinations of scenario sizes and pairwise modalities.

Table 3: GCMH Test Statistic of Pairwise Treatment Comparison

Pairwise Comparison	2×2	3×3	4×4	5×5
Audio | Visual	16.3	63	42.2	55.1
Audio | Audio & Visual	10.7	37.3	54.5	6.19*†
Visual | Audio & Visual	34.4	68.3	64.3	62.4

*Not a statistically significant difference

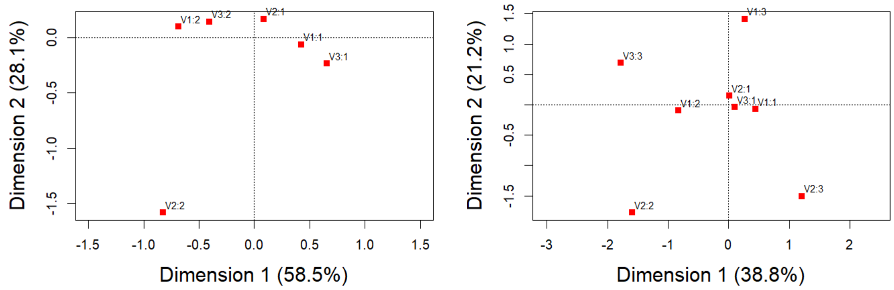
† Singular covariance matrix; Moore-Penrose psuedoinverse utilized

Additional insights regarding the players' behavior result from comparing results across the rows and columns in Table 3. The presentation modalities induced statistically distinct behavior in each of the 2×2, 3×3, and 4×4 games. Such a result implies the relevance of the multimedia learning tenets and that these principles are applicable over myriad complexity levels. However, Table 3 also yields a counter-intuitive result: behavior was not statistically different between the audio-only and the audio-and-visual modalities in the 5×5 games.

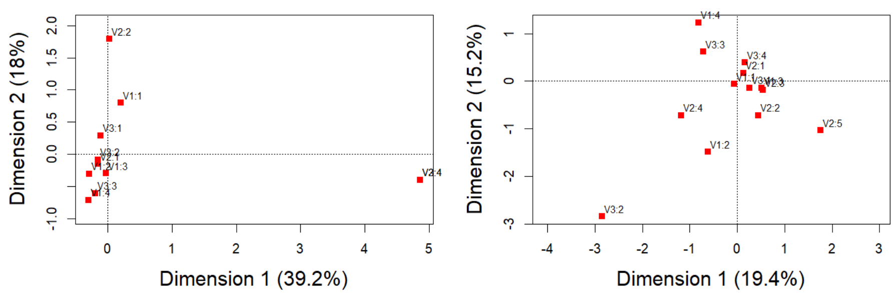
Although our experimentation does not permit the identification of a causative factor, multiple conjectures are logically plausible. Foremost, is the possibility that the Moore-Penrose pseudoinverse induced an overly conservative test statistic in this analysis. Additionally, test fatigue is another possible factor; the audio-and-visual modality of size 5×5 was the final question in this block of twelve questions. Finally, it is also plausible that participants were sufficiently overloaded with information for both the audio and audio-and-visual modalities of the largest-size game so as to reduce decision-making quality comparably towards level-0 thinking. This final conjecture supports the idea that, for even the relatively small 5×5 games, effective presentation of information may not suffice; additional decision-making support (e.g., an expert system) may be necessary.

Raw Response Analysis: Identifying Associations The MCA conducted herein implies that presentation modality affected each participant's individual behavior in a systematic manner. Figure 1 presents the MCA plots in principal coordinates by scenario size as calculated with the Burt matrix. Each point represents a possible response to a question, and each point has two numbers in its associated label. The first number is the presentation modality (i.e., 1 for audio-only, 2 for visual-only, 3 for audio-and-visual), and the second number indicates an action selected in the game. Actions never chosen by the participants are not depicted in the corresponding plot. For the 2×2, 3×3, 4×4, and 5×5 games, note that the total inertia is 86.6%, 60.0%, 57.2%, and 34.6%, respectively.

Utilizing the MCA plots, we can analyze each dimension to determine what factors best describe the behavioral variability. The horizontal dimensions of Figures 1a-1d can respectively be interpreted as separating players based on their selections across modalities. Scenarios 2×2 and 3×3 respectively exhibit a separation of Actions 1 and 2 and a separation of Actions 1 and 3 from Action 2. The horizontal dimension for Figure 1c shows a separation in Action 4 for the visual-only and audio-and-visual modalities. Similarly, Figure 1d delineates between actions in the audio-and-visual modality. The vertical dimension of Figures 1a, 1b, 1c, and 1d can be interpreted, respectively, as separating players who played Action 2 in the visual-only modality for sizes 2×2 and 4×4; differentiating the selection of Action 3 in the audio-only modality for size 3×3; and selection of Action 2 in the



(a) 2 × 2 Scenarios (b) 3 × 3 Scenarios



(c) 4 × 4 Scenarios (d) 5 × 5 Scenarios

Figure 1: Two-dimensional MCA Projections by Scenario Size

audio-and-visual modality for size 5 × 5.

Furthermore, the MCA plots can also be leveraged to identify subsets of players that tended to behave in a similar or dissimilar manner. In Figure 1a, notice the proximity between Points 1:2 and 3:2 as well as between Points 1:1 and 3:1. This implies that participants who chose the dominant (dominated) action in the audio-only modality tended to select the dominant (dominated) action in the audio-and-visual modality; however, this association did not translate to the visual-only modality. Figure 1b indicates negative associations were present based upon player selections in the visual-only modality; selecting Action 3 in the visual-only modality suggested a tendency to not select Action 3 in the audio-and-visual modality, whereas selecting Action 2 in the visual modality implies a tendency to not select Action 3 in the audio-only modality. Figure 1c depicts a similar negative association between the selection of Action 2 in the visual-only modality and Action 4 in the audio-and-visual modality, whereas a positive association can be observed between Action 4 selections within the same modalities. Figure 1d illustrates another negative association between selecting Action 2 in the visual-only and audio-and-visual modalities. Collectively, these results imply individual behavior varied systematically by presentation modality and suggest distinct steps of strategic reasoning were induced.

Cognitive-Hierarchy Analysis: τ-value Estimation Table 4 presents the participants' estimated τ-value for each question. Notable for each game size is the increase of τ from an audio-only to a visual-only presentation. Increases also exist between the visual-only and audio-and-visual modalities for the 3×3 and 4×4 games. The largest MSE between the participants' empirical distribution and the predicted CH distribution was 0.05 from the audio-only question of size 5×5. Half of all questions yielded an MSE less than 0.005.

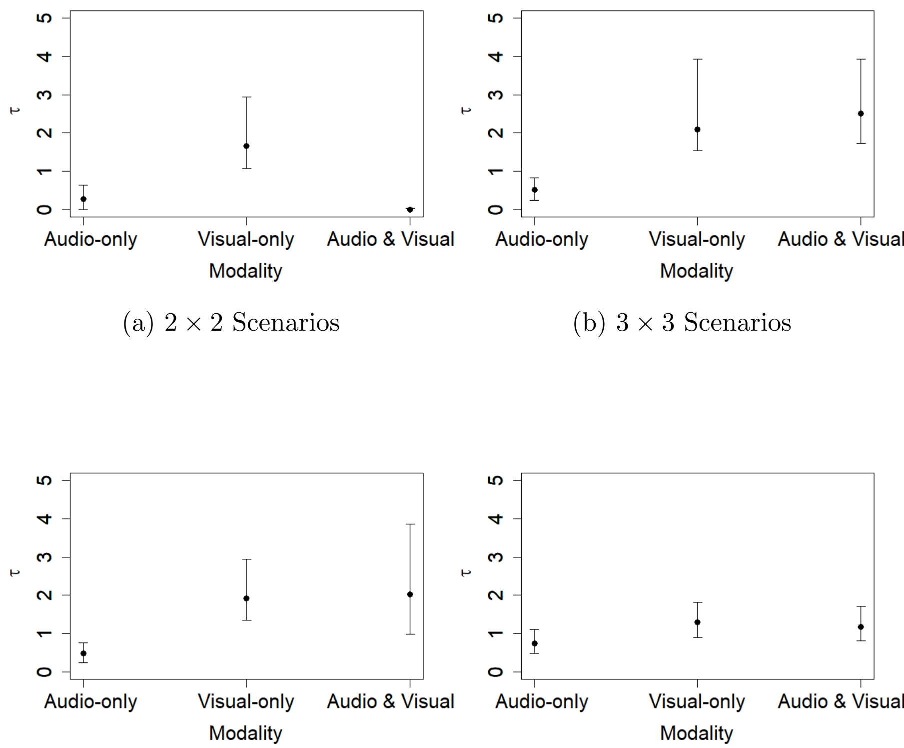
Table 4: Estimated τ-values for All Questions

Presentation Modality	2×2	3×3	4×4	5×5
Audio	0.27	0.52	0.48	0.74
Visual	1.65	2.09	1.92	1.29
Audio & Visual	0.00	2.50	2.02	1.18

Generally speaking, these point estimates suggest that the visual-only presentation induces higher steps of strategic thought across all complexity levels than the audio-only presentation. However, the results are less clear when moving from visual to audio-and-visual presentations. In the 3×3 and 4×4 games, the τ-values increased in accordance with the associated multimedia learning tenets, but results from the 2×2 and 5×5 games do not. It appears that the value of effective information presentation is not as significant as the problem size with 5×5 games, reinforcing the conjecture about the limits of effective communication. However, since the values in Table 4 are point estimates, we cannot yet make statistically significant statements about such differences; to do so, we must leverage our bootstrap confidence intervals.

Conditioned on the presentation modality, we generated 10,000 resamples of bootstrap distributions of τ-estimates from the observed data. When all resampled answers were the dominant action, τ was estimated to be 5.00 (i.e, the upper bound on the preset range) because the best fitting τ to the observed selection approaches infinity. For each game size, the bootstrap distribution's median for the visual-only question was greater than the corresponding median for the audio-only question. Question sizes of 3×3 and 4×4 show the audio-and-visual questions yielding a generally higher τ-value relative to audio-only questions; however, there is a less pronounced distinction from the visual-only questions. Likewise, for these games, τ-estimates in the audio-and-visual modality appear more variable than other presentations. Scenario sizes of 5×5 have the most considerable overlap in the τ estimates, showing less positive influence resulting from the additional presentation modality, as was the case for the 3×3 and 4×4 scenarios. The bootstrap distributions in the 2×2 games would follow the tenets of multimedia learning, except that the audio-and-visual modality's distribution is tightly clustered near zero. We discuss this anomaly in more detail subsequently.

Figure 2 displays the percentile confidence intervals resulting from the bootstrap distributions; a more in-depth review of these confidence intervals provides additional perspective about the population's behavior. In the 2×2, 3×3, and 4×4 games, the intervals do not overlap, implying statistically different τ-values. Conversely, whereas such an overlap does exist in the 5×5 games, the upper bound of the visual-only confidence interval exceeds the upper bound of the audio-only modality, indicating greater levels of strategic thought are more likely to be observed in the former. Similarly, the audio-and-visual modality's upper bound exceeds the upper bound of the audio-only modality in all cases but the anomalous 2×2 scenario.



(a) 2 × 2 Scenarios

(b) 3 × 3 Scenarios

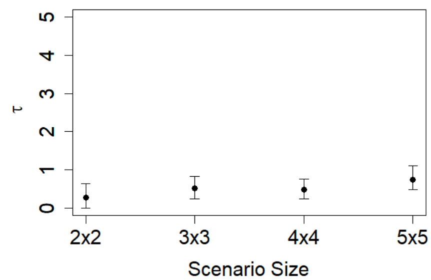
(c) 4 × 4 Scenarios

(d) 5 × 5 Scenarios

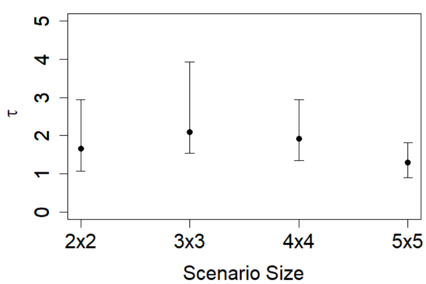
Figure 2: Percentile Confidence Intervals Across Scenario Size

As presented in Figure 3, similar confidence intervals allow one to examine the effect of situational complexity on reasoning depth. Although few statistical differences were observed, these intervals provide interesting insight. In general, the overlap suggest that the induced reasoning depth is relatively similar within a presentation modality. Moreover, it is interesting to note that, in the audio-only presentation, every confidence interval is almost entirely contained in [0,1]. This result comports with multimedia learning's limited capacity assumption and, when compared against the other modality's confidence intervals, suggests player's leveraged less information received in this modality. Alternatively, the visual-only modality induced higher τ-estimates, but these were relatively similar to those identified in the audio-and-visual modality. From 3×3 games to larger sizes in the visual-only and audio-and-visual modalities, the upper bounds on confidence intervals is decreasing, with a notable drop off in the 5×5 games. Thus, the positive effect of leveraging multimedia principles has readily discernible limitations on improving strategic decisions in more complex situations.

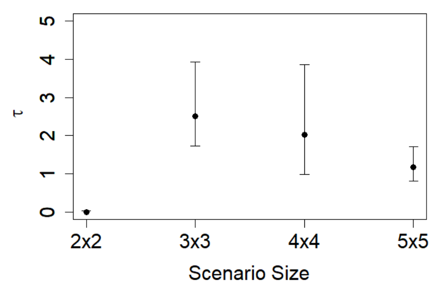
Collectively, this analysis explains all the differences observed in Table 4 except for the anomalous behavior in the 2×2 audio-and-visual setting. Based on a *post hoc* analysis of the test question and participant feedback, it appears that an unforeseen covariate affected their behavior. The game in question was a modified Prisoner's Dilemma wherein the player's best response was to "lie". Participant feedback suggested that, although many participants recognized it as the utility-maximizing-action, they *did not view lying as an action they could willingly select*. Others believed



(a) Audio-only



(b) Visual-only



(c) Audio-and-Visual

Figure 3: Percentile Confidence Intervals Across Modality

their opponent shared this ethical framework and would not lie either, thereby making honesty the preferred action. Likely relevant to the seemingly non-strategic group performance is that every participant was either an active duty military or civilian employee in the Department of Defense, taking classes at the Air Force Institute of Technology, a technical graduate school run by the United States Air Force. Among the three core values of the Air Force, the first value is "Integrity First" (US Air Force 2015); honesty is valued above all else. Such a principle-based, voluntary deviation from utility maximizing behavior would explain this anomaly, and this anecdotal inference invites further study beyond the scope of this research.

Implications and Limitations The results described herein illustrate the relevance of multimedia learning to establish a guiding set of best communication practices founded in multimedia research. It has been shown that alternative presentation modalities induce statistical differences in decision maker behavior. Moreover, such differences were shown to vary systematically. These systematic differences can be linked to increased (decreased) levels of strategic reasoning. It was shown that audio-centric communications are unlikely to promote sophisticated decision-making, even in the simplest competitive interactions. Alternatively, our results suggest that the communication of information must ensure visual information is provided to a decision maker to promote high-quality decisions. The same communication could also adopt a combined audio-and-visual approach; however, the marginal benefits of doing so are relatively muted. Such patterns were observed across varying levels of situational complexity implying that, even for the simplest competitive environments, the aforementioned implications remain relevant. Our results suggest audio-and-visual presentation invoke diminishing returns as situational complexity increases, indicating a threshold beyond which decision makers need an expert system to inform decisions. Finally, it was shown that the strategic-reasoning depths induced by a presentation modality are relatively similar across situational complexities, thereby implying the relevance of these design implications across multitudinous competitive environments.

CONCLUSION

The empirical data from these experiments was analyzed utilizing multiple statistical techniques, both without and with game-theoretic assumptions. Namely, the statistical effect of presentation modality on the participant's behavior was first examined via the GCMH test and, if the behavior was deemed distinct, MCA was performed to identify associations across participant responses. Subsequently, game-theoretic analysis was performed under CH assumptions. For each game, the average steps of thought utilized by the participants was estimated, and bootstrapped confidence intervals about the values were obtained. Hypothesis tests leveraging the confidence intervals were subsequently evaluated to determine the effect of presentation modality and situational complexity on the participants' strategic abilities. Based on this statistical testing, it was found that the tenets of multimedia learning do generally extend effective communication of information to decision-makers, but not uniformly across situational complexities. Moreover, these principles have limited benefit to inform strategic decisions for relatively more complex scenarios.

References

Agresti, A. (2002), *Categorical Data Analysis*, John Wiley & Sons, New York, NY.

Alsaleh, R. & Sayed, T. (2022), 'Do road users play Nash equilibrium? a comparison between nash and logistic stochastic equilibriums for multiagent modeling of road user interactions in shared spaces', *Expert Systems with Applications* **205**, 117710.

Arad, A., Grubiak, K. P. & Penczynski, S. P. (2022), 'Does communicating within a team influence individuals' reasoning and decisions?', *Experimental Economics* **2022**, 1–21.

Camerer, C. F. (2011), *Behavioral Game Theory: Experiments in Strategic Interaction*, Princeton University Press, Princeton, NJ.

Camerer, C. F., Ho, T. H. & Chong, J. K. (2004), 'A cognitive hierarchy model of games', *The Quarterly Journal of Economics* **119**(3), 861–898.

Clark, J. M. & Paivio, A. (1991), 'Dual coding theory and education', *Educational psychology review* **3**(3), 149–210.

DeLong, K. L., Syrengelas, K. G., Grebitus, C. & Nayga Jr, R. M. (2021), 'Visual versus text attribute representation in choice experiments', *Journal of Behavioral and Experimental Economics* **94**, 101729.

Efron, B. & Tibshirani, R. J. (1994), *An Introduction to the Bootstrap*, CRC press, Boca Raton, FL.

Greenacre, M. & Blasius, J. (2006), *Multiple Correspondence Analysis and Related Methods*, CRC press, Boca Raton, FL.

Holm, H. J., Nee, V. & Opper, S. (2020), 'Strategic decisions: behavioral differences between CEOs and others', *Experimental Economics* **23**(1), 154–180.

Li, C., Turmunkh, U. & Wakker, P. P. (2019), 'Trust as a decision under ambiguity', *Experimental Economics* **22**(1), 51–75.

Mayer, R. E. (2009), *Multimedia Learning*, Cambridge University Press, Cambridge, UK.

Mayer, R. E. (2010), 'Applying the science of learning to medical education', *Medical Education* **44**(6), 543–549.

Paivio, A. (1969), 'Mental imagery in associative learning and memory', *Psychological Review* **76**(3), 241–263.

Serra, D. (2021), 'Decision-making: from neuroscience to neuroeconomics—an overview', *Theory and Decision* **91**(1), 1–80.

Shoham, Y. & Leyton-Brown, K. (2008), *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*, Cambridge University Press, Cambridge, MA.

Stahl, D. O. & Wilson, P. W. (1995), 'On players' models of other players: Theory and experimental evidence', *Games and Economic Behavior* **10**(1), 218–254.

Terzopoulou, Z. & Endriss, U. (2022), 'Strategic manipulation in judgment aggregation under higher-level reasoning', *Theory and Decision* **92**(2), 363–385.

US Air Force (2015), *Air Force Doctrine Document 2 - Leadership*, LeMay Center for Doctrine, Maxwell Air Force Base, AL.
URL: *https://www3.nd.edu/~jthomp19/AS300/1_Fall%20Semester/Air_Force_Leadership/Supplemental_Reading/Air_Force_Doctrine_Vol_II_Leadership.pdf*

Vaid, J. (1988), 'Bilingual memory representation: A further test of dual coding theory.', *Canadian Journal of Psychology/Revue canadienne de psychologie* **42**(1), 84.