mathematics

MDPI

Article

A Context-Sensitive Alternative to Hick's Law of Choice Reaction Times: A Mathematical and Computational Unification of Conceptual Complexity and Choice Behavior

Ronaldo Vigo [1,2,*], Charles A. Doan [2,3], Jay Wimsatt [2,4] and Cody B. Ross [1,2]

1 Department of Psychology, Ohio University, Athens, OH 45701, USA; cr355120@ohio.edu
2 Consortium for the Advancement of Cognitive Science, Ohio University, Athens, OH 45701, USA
3 Department of Psychology, Marietta College, Marietta, OH 45750, USA; cad007@marietta.edu
4 Department of Psychology, Arizona State University, Tempe, AZ 85281, USA; jay.wimsatt@asu.edu
* Correspondence: vigo@ohio.edu; Tel.: +1-740-593-1707

Abstract: Hick's law describes the time that individuals take to make a preference decision when presented with a set of possible choices. Basically speaking, the law states that decision time is a logarithmic function of the number of choices when the choices are equiprobable. However, the evidence examined here suggests that this, and a variant of the law for non-equiprobable choices based on Shannon entropy, are not effective at predicting decision reaction times involving structured sets of alternatives. The purpose of this report is to communicate a theoretical alternative to Hick's law that is derived from a mathematical law of invariance for conceptual behavior at the heart of Generalized Invariance Structure Theory (Vigo, 2013, 2015). We argue that such an alternative accounts more precisely for decision reaction times on structured sets. Furthermore, we argue that Hick's law is a special case of this more general law of choice reaction times for categories with zero degree of invariance.

Keywords: Hick's law; choice law; invariance structure theory; computational cognitive model; choice response times; categorical invariance; subjective complexity

MSC: 91E10; 91E30; 91E40; 91E45

Citation: Vigo, R.; Doan, C.A.; Wimsatt, J.; Ross, C.B. A Context-Sensitive Alternative to Hick's Law of Choice Reaction Times: A Mathematical and Computational Unification of Conceptual Complexity and Choice Behavior. *Mathematics* **2023**, *11*, 2422. https://doi.org/10.3390/math11112422

Academic Editor: Takashi Yamauchi

Received: 2 April 2023
Revised: 12 May 2023
Accepted: 18 May 2023
Published: 23 May 2023

1. A Context-Sensitive Alternative to Hick's Law of Choice Reaction Times: Unifying Conceptual Complexity and Choice Behavior

Given a set of n alternatives, Hick's law describes the average time it takes for individuals to select any one alternative from the set. The simplest form of Hick's law [1] states that, given n equiprobable choices, the average reaction time T required to choose among the choices is proportional to the logarithm (base two) of the number of choices n, as shown in Equation (1) below, where b is a constant that can be determined empirically by fitting a line to measured data.

$$T = b \cdot \log_2(n+1) \tag{1}$$

It was argued by Hick that his law has a base two logarithmic form because individuals may be systematically subdividing the total set of alternatives into a pair of categories at a time. This facilitates eliminating, at each step, one half of the alternatives rather than requiring the linear-time process of considering each alternative at a time. Thus, the logarithmic function may be interpreted as describing a depth of a choice tree hierarchy with a sub-decision at each step that implies that a binary search is performed. The addition of 1 to the value n does the following: (1) resolves the undefined case of a set without alternatives (i.e., accounting for the undefined case of $\log_2(\cdot)$ of zero), and (2) regards the possibility of no stimulus signal as an alternative, thereby raising the number of alternatives to $n + 1$ alternatives.

A more general version of the law, referred to as Equation (2) below, is more directly related to Shannon entropy [2,3] in that it predicts reaction times for the case of choices with unequal probabilities. The law can be generalized as follows, where again b is a constant that can be determined empirically by fitting a line to measured data and p_i refers to the probability of the i-th alternative yielding the information-theoretic entropy (for details, see Hick's original article [1]). Note that, when $p_i = 1/n$, then Equation (2) reduces to the basic form of the law described by Equation (1).

$$T = b \cdot \sum_{i}^{n} p_i \log_2\left(\frac{1}{p_i} + 1\right) \tag{2}$$

As mentioned, Hick's law has a base two logarithmic form because the presupposed processing strategy for selecting an item assumes that people subdivide the total collection of choices into two categories at a time, eliminating about half of the remaining choices at each level of analysis. As such, the logarithmic function then reveals the number of levels of analysis required for a particular choice set. However, this manner of processing information does not consider the relationship between the features or characteristics of the items in the choice set, especially if these are partially shared. Furthermore, it was this very deficit that prompted criticism toward the use of Shannon information theory in psychological research. For example, Luce [4] points out that:

"However, in my opinion, the most important answer lies in the following incompatibility between psychology and information theory. The elements of choice in information theory are absolutely neutral and lack any internal structure; the probabilities are on a pure, unstructured set whose elements are functionally interchangeable" (p. 185).

Correspondingly, Donald Laming, whose work in *Information Theory of Choice-Reaction Times* [5] represents one of the most ardent attempts at linking information theory and choice-reaction times, later stated: "This idea does not work. While my own data [5] might suggest otherwise, there are further unpublished results that show it to be hopeless" [6], pp. 642. Despite these admissions, Hick's law has been applied to experimental and applied cognitive research for decades. Examples include applications to intelligence testing [7], ergonomics research [8,9], and human–computer interfaces [10–12] (also see Liu et al. [13] for a contrary view). See Proctor and Schneider [14] for a comprehensive review of the history of Hick's law, contemporary models, and additional applications.

In previous works by the first author, entropy-based accounts of psychological phenomena are disputed on similar theoretical grounds to those discussed by Luce [15–18]. More recently, Vigo et al. [18] demonstrated how a theory of concept learning difficulty derived directly from Shannon entropy [19] does not account for key results in human conceptual behavior (for a gentle introduction to concepts, see [20]). Here, we extend these arguments to demonstrate the weaknesses of Hick's law with respect to structured sets of object stimuli. Importantly, it is not our aim to simply communicate the shortcomings of Hick's law; indeed, Kveraga et al. [21] and Pavão et al. [22], among others [23,24], have already identified weaknesses with regard to saccades and sequential learning, respectively. Instead, the aim is to show how an alternative and more general law derived from a theory of category learning and conceptual behavior can overcome these shortcomings with respect to the general domain of categories (i.e., structured choice sets). While Pavão et al. [22] introduced a joint entropy model as an alternative to Hick's law, it was intended explicitly for sequence learning situations and not for structured sets. As far as we know, this is the first time that a direct mathematical link between concept learning and decision-making that successfully accounts for reaction time data has been achieved.

The next two sections contain a brief description of the theory and a simple derivation of the alternative law. To reach as wide a mathematical readership as possible, this manuscript is written in a mathematical/computational modeling style, which is typical in the fields of applied sciences, applied mathematics, and engineering mathematics. In other words, derivations are conducted in a natural straightforward manner rather than using the axiomatic method. Empirical evidence will be used to assess the effectiveness and veracity of the mathematical candidate laws derived and tested. For more detailed and/or more formal accounts, the reader is referred to other works by the first author [15–17,25–27] and the technical appendix located in the Supplementary Materials.

2. GIST and the Law of Invariance

One of the ultimate goals of human categorization research is to determine how difficult individuals find it to learn different types of categories where a category is construed as a set of entities with shared dimensions whose dimensional values (i.e., features) are related in some way. (Whenever all the elements of a category share the same dimensions, the category is referred to as regular. Whenever different elements share different sets of dimensions, the category is referred to as irregular. In this paper we focus only on regular categories.) Accordingly, we also refer to categories as *structured sets*. The law of invariance (LOI; [16,17]), a product of *Generalized Invariance Structure Theory* (GIST; [16,17]), was designed to make such predictions. Its effectiveness stems from the ability to account for how humans acquire relational information from the object stimuli of categories, or in other words, how humans assess the objective relationships that may exist between the objects that make up categories. A simple example of the relationships that may exist between objects of a category that is dimensionally defined comes from considering the space of objects that may be generated by three binary dimensions (e.g., *shape*, *size*, and *color*). Such a space consists of 2^D objects where D is the number of binary dimensions (which, in this case, is three, each with two possible feature values). Subsets from the eight objects may then be selected to form categories whose object features are related in particular ways. For the case of categories containing four objects, there are only six possible category structures [28]. Examples of such structures are displayed in Figure 1.

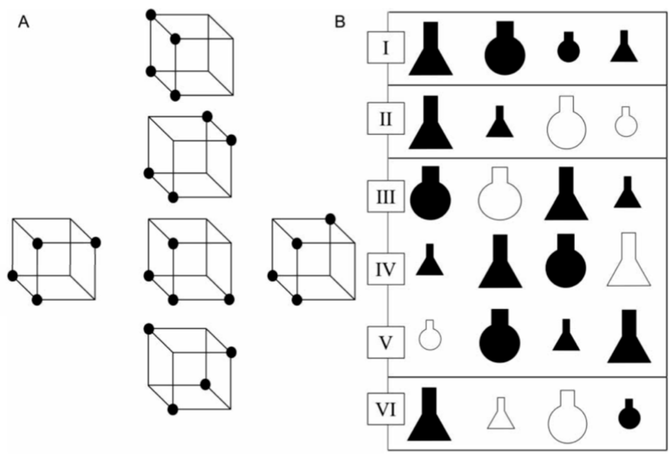

Figure 1. Boolean Cube (**A**) and Stimulus Representation (**B**) of the six logically distinct $3_2[4]$ structure types. Each dot on the Boolean Cube (panel **A**) represents a member of the four-object categorical stimulus. The three binary dimensions for the stimulus representations in panel (**B**) are *shape* (triangular, circular), *size* (small, large), and *color* (white, black). Adapted from "The GIST of Concepts", by R. Vigo, 2013, *Cognition*, *129*, p. 142. Copyright 2013 by Elsevier B.V.

Specifically, the LOI, as shown in Equation (3) below, asserts that the subjective degree of concept learning difficulty ψ of a category of objects X (as measured by the error rates in classifying its members) is *directly proportional to the cardinality of the category* (i.e., number of its elements $|X|$) and *inversely proportional to the square of the exponent of the degree of categorical invariance* $\widehat{\Phi}(X)$ of the category X, where $\widehat{\Phi}^2(X)$ stands for $\left(\widehat{\Phi}(X)\right)^2$. In a nutshell, the degree of categorical invariance $\widehat{\Phi}(X)$ of a category X is obtained by first applying the structural manifold operator Λ (defined in the technical appendix located in the Supplementary Materials to the article) to X, and then applying the Euclidean metric to determine the psychological distance from the computed structural manifold $\Lambda(\mathrm{X})$ to the zero structural manifold. More specifically, the Λ operator extracts invariance pattern information from the category as a vector of the proportion of perceived dimensional symmetries within (i.e., as a structural manifold). It does this via a core cognitive mechanism referred to as "dimensional binding" that is proposed to underlie all human conceptual behavior [16,17]. The mechanism basically captures a hypothesized process involving the systematic suppression of dimensions via rapid attention shifting while assessing degrees of partial similarity. This process captures the fundamental role that relational cognition plays in category learning and the types of relevant relations (i.e., invariance patterns) between object stimuli that overcome the limitations encountered by Shannon information as described by Luce [4].

$$\psi(\mathrm{X}) = |\mathrm{X}| \exp\left(-k\widehat{\Phi}^{2}(\mathrm{X})\right) \tag{3}$$

The scaling parameter $k \geq 0$ (where k is a real number) plays the role of a structure discrimination parameter that, in the version without free parameters shown in Equation (4), is substituted by the index D_0/D to characterize this discrimination capacity. In the index, D_0 stands for the minimum number of dimensions needed to non-trivially describe a category of objects, namely two, and $D \geq 2$ is the number of dimensions that define the categorical stimulus. More specifically, the parameter k indicates the overall ability of an observer to discriminate structures from a zero-invariance structure as a function of the number of dimensions that define the category, thus summarizing concept learning performance. This parameter is intended to capture the ability of an observer to extract invariance patterns at different dimensional levels of analysis, with higher values denoting superior discrimination. It should be acknowledged that the parameter k does not convey much useful information unless the overall law supplies fairly accurate fits to the data. The single free parameter in this parametric version also makes it possible to account for individual differences in concept learning capacity and performance between humans.

Results from experiments conducted by several researchers, including the authors, have corroborated both variants of the law. For example, the LOI without free parameters and with one free-parameter accounts for about 90% of the variance in data from 84 types of category structures sampled from 5100 distinct categories [16,17]. In addition, the LOI has accounted for over 90% of the variance in results from experiments by Shepard et al. [28], Nosofsky and Palmeri [29], Vigo and colleagues [18,30,31], and many others.

$$\psi(\mathrm{X}) = |\mathrm{X}| \exp\left(-\left[\frac{D_0}{D}\right]\widehat{\Phi}^{2}(\mathrm{X})\right) \tag{4}$$

3. Derivation of the Invariance-Based Choice Reaction Times Law

The conceptual-choice principle, introduced by Vigo [17], allows us to quantitatively define a relationship between choice behavior and concept learning: *for any two finite choice sets X and Y consisting of the same number of object stimuli defined over the same set of dimensions, if the degree of concept learning difficulty of X $\geq$ Y, then the degree of choice difficulty C_d in choosing a preferred item* r *from X is less than or equal to the degree of choice difficulty C_d in choosing a preferred item* s *from Y*. This is stated formally as follows:$C_d(r|\mathrm{X}) \leq C_d(s|\mathrm{Y})$. We can infer from this principle that, under the same assumptions, the amount of time required to choose a preferred item r from X is less than or equal to the time it takes to choose a preferred item s from Y; that is, $C_{RT}(r|\mathrm{X}) \leq C_{RT}(s|\mathrm{Y})$. More generally see Equation (5),

$$C_{RT}(r\,|\mathrm{X}) \propto \frac{1}{\psi(\mathrm{X})} \tag{5}$$

which predicts that the choice reaction time (C_{RT}) associated with any preferred item r being chosen from any choice set X is inversely proportional to the subjective degree of concept learning difficulty of X (note that capital letters of the English alphabet are used to represent arbitrary dimensionally defined sets of objects or, equivalently, categorical stimuli; for more formal details, see the technical appendix located in the Supplementary Materials to the article). This naturally follows because, as shown in Equation (3), the learnability of a concept increases as the degree of categorical invariance of the category from which it is learned increases. However, essentially, categorical invariance is a measure of the degree of perceived logical coherence of a category. Thus, although categorical coherence has a beneficial side in that it makes it easier to learn and grasp concepts by facilitating the assessment of object stimuli interactions and featural/dimensional level diagnosticity, it also increases the likelihood of confusing object stimuli in the category with one another. Consequently, for structured choice categories defined on the same dimensional space, as is the case above, choices from those that are relatively more difficult to learn, on average, take less time due to a corresponding reduction in the object confusability resulting from their lower coherence.

Although the inverse $1/\psi(\mathrm{X})$ adequately approximates choice reaction times, according to Vigo [17], the influence of category cardinality (i.e., the number of items in the category) is diminished by taking the inverse of ψ, and it is therefore not the most accurate characterization (that is, it does not yield the most accurate and most meaningful predictions). To obtain a more meaningful and more accurate characterization, the inverse of the degree of categorical invariance and the cardinality of the set must both significantly influence choice RTs. However, the ideal approach is not to formulate choice RTs as directly proportional to the inverse of the degree of categorical invariance of a choice set and to its cardinality because the aim is to formulate an expression in terms of ψ (recall that the original objective per Vigo [17] was to establish a direct connection between concept learning via the law of invariance ψ and decision-making reaction times).

Based on this reasoning, Vigo [17] devised the desired precise quantitative connection between the inverse of degree of concept learning difficulty and choice reaction time by formulating Equation (6) (found below) in terms of degree of concept learning difficulty. The equation states that choice reaction times are determined by the sum of the square of the degree of categorical invariance (also known as degree of logical coherence or logical homogeneity) of X and the natural logarithm of the cardinality of X. Taking the natural logarithm reveals the precise additive relationship sought between the influence of structure as measured by degree of invariance $\widehat{\Phi}$, set "size" $|\mathrm{X}|$ (where $|\mathrm{X}|$ stands for the cardinality or number of items in X), and choice RT.

$$C_{RT}(r\,|\mathrm{X}) \propto \log_e\left(\frac{|\mathrm{X}|^2}{\psi(\mathrm{X})}\right) = \log_e\left(\frac{|\mathrm{X}|^2}{|\mathrm{X}|e^{-k\widehat{\Phi}^2(\mathrm{X})}}\right) = \log_e\left(|\mathrm{X}|e^{k\widehat{\Phi}^2(\mathrm{X})}\right) = k\widehat{\Phi}^2(\mathrm{X}) + \log_e(|\mathrm{X}|) \tag{6}$$

Equation (6) above can be reformulated without free parameters as follows:

$$C_{RT}(r\,|\mathrm{X}) \propto \log_e\left(\frac{|\mathrm{X}|^2}{\psi(\mathrm{X})}\right) = \log_e\left(\frac{|\mathrm{X}|^2}{|\mathrm{X}|e^{-\left(\frac{D_0}{D}\right)\widehat{\Phi}^2(\mathrm{X})}}\right) = \log_e\left(|\mathrm{X}|e^{\left(\frac{D_0}{D}\right)\widehat{\Phi}^2(\mathrm{X})}\right) = \left(\frac{D_0}{D}\right)\widehat{\Phi}^2(\mathrm{X}) + \log_e(|\mathrm{X}|) \tag{7}$$

Note that, in Equations (6) and (7), when the degree of categorical invariance $\widehat{\Phi}$ of the category in question is zero, the equation reduces to the simple variant of Hick's law but features e as the base of the logarithm instead. Furthermore, note that the above equations assume that the categories are non-empty (i.e., $\mathrm{X} \neq \phi$) and that the number of dimensions attributable to the set of items is at least two (i.e., $D \geq 2$). (However, to account for empty categories, one could modify the derived final expressions in Equations (6) and (7) above by replacing $|X|$ with $|X| + 1$) Equation (6) is henceforth referred to as the Invariance-based Choice RT law (ICRT), and Equation (7) is its counterpart without free parameters (ICRT-NP).

4. Empirical Support

Empirically, the present study tests and compares the two variants of Hick's law, as shown in Equations (1) and (2), and the two variants of the ICRT. Both ICRT variants (with and without the single free parameter) are based on the inverse of the degree of concept learning difficulty ψ, and it is hypothesized in this manuscript, as in Vigo [17] and in Vigo and Doan [32], that increases in categorical invariance (i.e., internal coherence) among alternatives of a choice set increase the degree of choice difficulty and thus lead to longer choice reaction times. Furthermore, it is shown in this section that the ICRT, with its structure-sensitive basis, outperforms both variants of Hick's law. Data for our tests were derived from Vigo and Doan [32], a study where we presented to participants categories of object stimuli (i.e., categorical stimuli) conforming to different category structure types. Participants were then instructed to select the item they most preferred from each displayed categorical stimulus. Next, we describe the study in greater detail, along with model fit tests for both variants of Hick's law and both variants of the ICRT. Importantly, the cross-validation and bootstrapping analyses that we report represent new analysis methods as applied to the Vigo and Doan [32] data that help further differentiate the performance across the tested models.

As described above, Vigo and Doan [32] conducted a large-scale experiment whereby they compared reaction times for participants to select their preferred item from among a set of related items (categorical stimulus). Given that this 2015 experiment involved participants choosing objects from structured choice sets, we have extracted the RT data to test both iterations of Hick's law as well as the ICRT and its non-parametric variant. Each categorical stimulus varied in terms of its dimensionality (three- or four-dimensional structure) and the number of alternatives from which participants could select their preferred item (two, three, four, five, six, or twelve alternatives). In addition, the researchers assessed the robustness of preferential decisions for these category structure types by assessing three different stimulus sets (realistic clocks, shoes, and t-shirts; see Figure 2 for examples) and up to four different structure instances for each category structure per participant. Accordingly, the study allowed for a repeated measures statistical analysis and, by extension, less inherent variation and more than sufficient power for detecting differences between the structure types.

The first experiment involved assessing choice behavior for the well-known $3_2[4]$ structure types across 48 participants, with each structure consisting of four alternatives (denoted: [4]) defined over three binary-valued (denoted: 3_2) stimulus dimensions [16,25,28,33–36]. Due to their cognitive tractability, these structures have been repeatedly studied regarding their learning difficulty in categorization tasks, and this learning difficulty (in terms of increasing number of classification errors) is highly replicable across the six structure types (e.g., I < II < [III, IV, V] < VI) [28]. Regarding preferential decision making, participants made 10 choices per structure type and per stimulus type (only the shoe and t-shirt stimuli), resulting in analyses across ~960 choices per structure type.

The second experiment was more comprehensive and involved replicating the results of the first experiment with new participants and fewer trials per structure type (four trials; $N = 45$ participants) while also assessing choice behavior on 50 other logically distinct three- and four-dimensional category structures and a new stimulus type (realistic clocks). To accommodate the increase in tested structure types, we divided the fifty-six structure types into three groups of approximately equal size (E1 received the eighteen three-dimensional structure types, including the six $3_2[4]$ structure types; E2 received the nineteen $4_2[4]$ structure types; E3 received the nineteen $4_2[12]$ structure types). Each of these three groups of structure types involved assessing reaction times across a different sample of participants (N = 45, 44, and 41 for E1, E2, and E3, respectively). Next, we report first on the aggregate results for the six $3_2[4]$ structure types (study 1 + relevant data from E1) before discussing the results from the other two groups of participants (E2 and E3). Although we only report results from samples of moderate size, our extensive cross-validation and bootstrapping analyses suggest that the fits are stable across the three- and four-dimensional structure types.

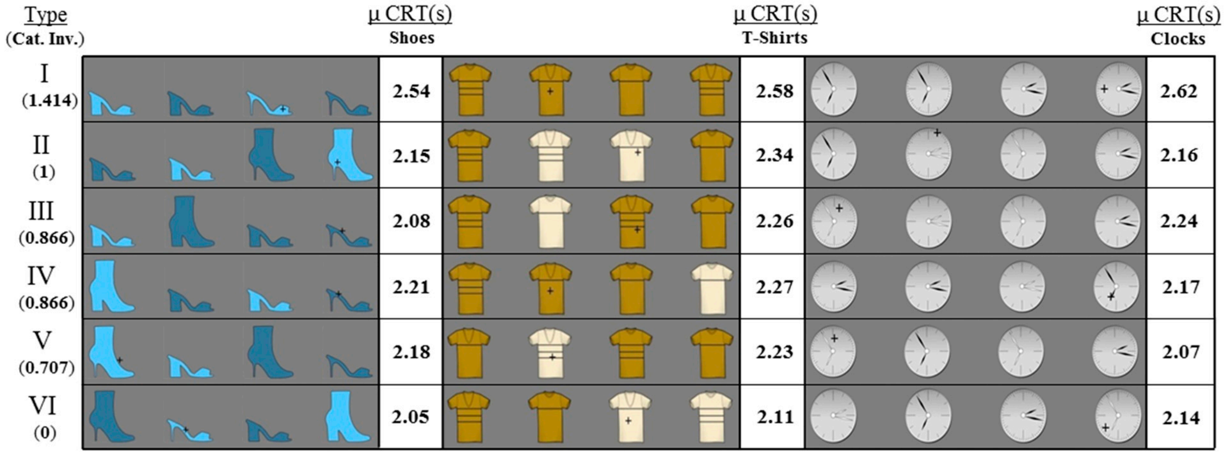

Figure 2. The $3_2[4]$ Structure Types Instantiated Across the Three Different Stimulus Sets Used by Vigo and Doan [32]. The leftmost column displays the degree of categorical invariance associated with each of the six structure types, whereas the average values provided thereafter refer to the average reaction times across participants for each of the six structure types for the first experiment conducted by Vigo and Doan [32]. The three stimulus dimensions for the shoes were (*size*: small, large; *color*: light blue, dark blue; *heel*: narrow, wide); the three stimulus dimensions for the t-shirts were (*color*: light beige, dark beige; *stripes*: one, three; *neck*: crew, v-neck); and the three stimulus dimensions for the clocks were (*hand color*: white, black; *hand angle*: acute, obtuse; *tick marks*: few, many). Adapted from "The Structure of Choice", by R. Vigo and C.A. Doan, 2015, *Cognitive Systems Research*, *36–37*, *p. 6*. Copyright 2013 by Elsevier B.V.

5. Reaction Times for the Six $3_2[4]$ Category Structures

Vigo and Doan [32] first investigated differences in choice reaction times for each of the six $3_2[4]$ category structure types. They extended upon robust categorization learning results by showing that they also vary in the time it takes participants to select their preferred item from such sets. Across 93 participants and ~9000 trials (~1500 trials per structure), they found that the structure types that were most difficult to learn from a conceptual standpoint (e.g., VI) are the same structure types from which it is easiest to select one's preferred item, whereas the opposite held true for the least difficult to learn from a conceptual standpoint (e.g., I and II). This pattern of choice reaction times held for all three stimulus types (clocks/shoes/t-shirts). The leftmost column of Figure 3 displays these differential reaction times for these six structure types aggregated across the three stimulus types, and how accurately Hick's law, Hick's entropy, and the ICRT-NP account for these reaction times. As can be seen, the ICRT-NP accounts for 97% of the variance in the $3_2[4]$ reaction times and outperforms both Hick's law and Hick's entropy, which account for only 0% and 20% of the variance respectively. (We calculated Hick's entropy for each category structure by first assessing the selection probabilities for all objects for each unique category instance. Then, we used Equation (2) and summed these probabilities across each instance to get the predicted time to select an object for that category instance. Because there were four tested category instances for each category structure, we then averaged these four times to get a final average predicted time for that category structure. This procedure was applied across the six category structures belonging to the $3_2[4]$ structure family and the nineteen category structures belonging to the $4_2[4]$ structure family.)

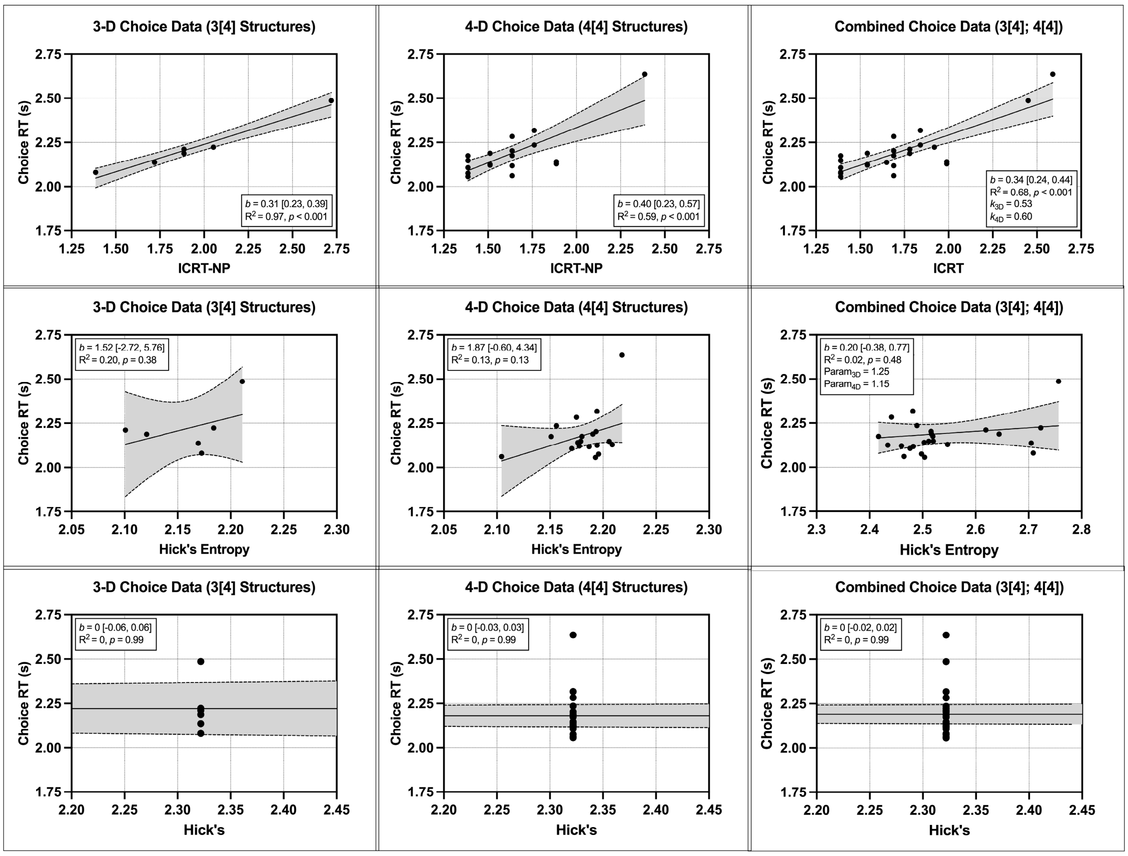

Figure 3. Linear regression fits to choice reaction times associated with the four-object three-dimensional, four-dimensional, and combined 3D/4D category structures assessed by Vigo and Doan [32]. The two *k*-parameters for the ICRT (rightmost plot, top row) and the two *b*-parameters for Hick's entropy (rightmost plot, middle row) were estimated using the solver add-in in Excel using the GRG non-linear solving method.

6. Reaction Times for the Nineteen $4_2[4]$ Category Structures

In addition to these six structure types, Vigo and Doan also assessed ~10,032 reaction times across 44 participants for the nineteen four-dimensional structure types that also consist of four alternatives (the $4_2[4]$ family of category structures; ~528 trials per structure). Similar to the three-dimensional results, they found that structure types that were harder and easier to learn were also the ones from which participants selected their preferred item in less and more time, respectively. These results, along with how well each model accounts for them, are shown in the middle column in Figure 3. The rightmost column of Figure 3 displays how well each model accounts for choice reaction times when both structure families are considered together ($3_2[4]$ and $4_2[4]$). Again, the ICRT-NP accounts for more of the reaction time variability among the four-dimensional structures (middle column; 59%), while the ICRT accounts for more of the reaction time variability when considering both the three- and four-dimensional four-object category structures ($3_2[4]$ and $4_2[4]$; 68%).

7. Reaction Times across 56 Logically Distinct Three- and Four-Dimensional Category Structures

Finally, Vigo and Doan extended their investigation beyond four-item choice sets by also presenting two-, three-, four-, six-, and twelve-item choice sets to participants (all were three-dimensional binary-valued structures except for the twelve-item sets). There were 12 logically distinct three-dimensional sets, and participants made a total of ~540 choices per set (~6480 total choices), divided evenly across the three separate stimulus types (clocks/shoes/t-shirts). For the 19 logically distinct four-dimensional sets, 41 participants made a total of ~492 choices per set (~9348 total choices). Consistent with the above results for the four-item sets, they again found converging evidence that sets which are more difficult to learn conceptually are easier to select from preferentially (and vice versa). Figure 4 displays how well each of the current models accounts for these additional structures and for the $3_2[4]$ and $4_2[4]$ structures, with the ICRT accounting for 87% of the variance in choice reaction times and Hick's law accounting for 76% of the variance in choice reaction times. The fact that the ICRT is able to achieve the fits displayed in Figures 3 and 4 without the incorporation of any free parameters demonstrates that a considerable proportion of the variation in choice RTs can be attributed to the core predictors of the model: the size of a choice set and its degree of internal coherence (i.e., invariance). Hick's law cannot account for the latter.

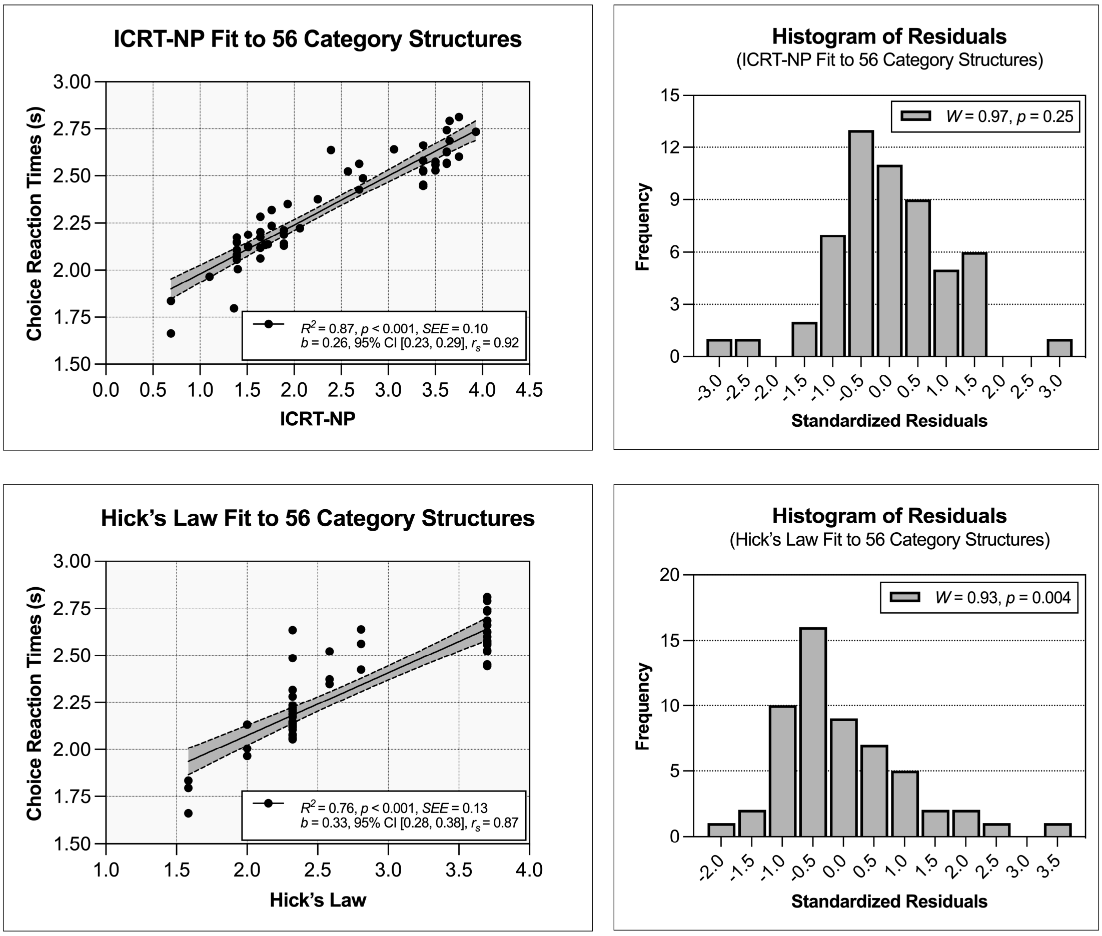

Figure 4. Linear regression fits of the ICRT and Hick's law to the choice reaction times associated with the 56 category structures (three- and four-dimensional) assessed by Vigo and Doan [32]. Model fit statistical results (e.g., R^2) are provided within each plot in the first column, while Shapiro–Wilk statistical results for the distribution of the standardized residuals are provided within each plot in the second column.

8. Bootstrapping Results

In addition to accounting for the average reaction times across the 56 category structures, we performed a bootstrapping analysis for the three- and four-dimensional structures that each consist of four alternatives (i.e., the six $3_2[4]$ structure types and the nineteen $4_2[4]$ structure types). This analysis was performed in R [37] and at three different levels, for 10,000 bootstrapped samples of sizes 15, 20, and 25 structures, where the latter assessment involved sampling the structures with replacement. Upon sampling each of 15, 20, and

25 structures, we assessed how well Hick's entropy and the ICRT-NP fit the sampled data by computing the R^2 and $RMSE$ for each linear model fit.

As shown in the top panel of Figure 5, the ICRT-NP accounts for a median value of 69% of the variance in reaction times for samples of size 15, whereas Hick's entropy accounts for 4% of the variance in reaction times. The middle and bottom panels reveal a similar pattern of results across both models, with the ICRT-NP accounting for 69% and 68% of the variance in reaction times for samples of size 20 and 25, respectively. Supplementing the linear fits, we also assessed the normality of the residual distributions for each of the 10,000 samples across the three different sample sizes. For the ICRT-NP, we found the median Shapiro–Wilk values (p-values) of 0.94 (0.40), 0.95 (0.39), and 0.92 (0.11), whereas, for Hick's entropy, we found the median Shapiro–Wilk values (p-values) of 0.81 (0.005), 0.78 (<0.001), and 0.83 (0.002). In sum, the ICRT-NP generally provides a more accurate and stable fit of the data across all three bootstrapped sample sizes when compared to Hick's entropy-based law for non-equiprobable choices.

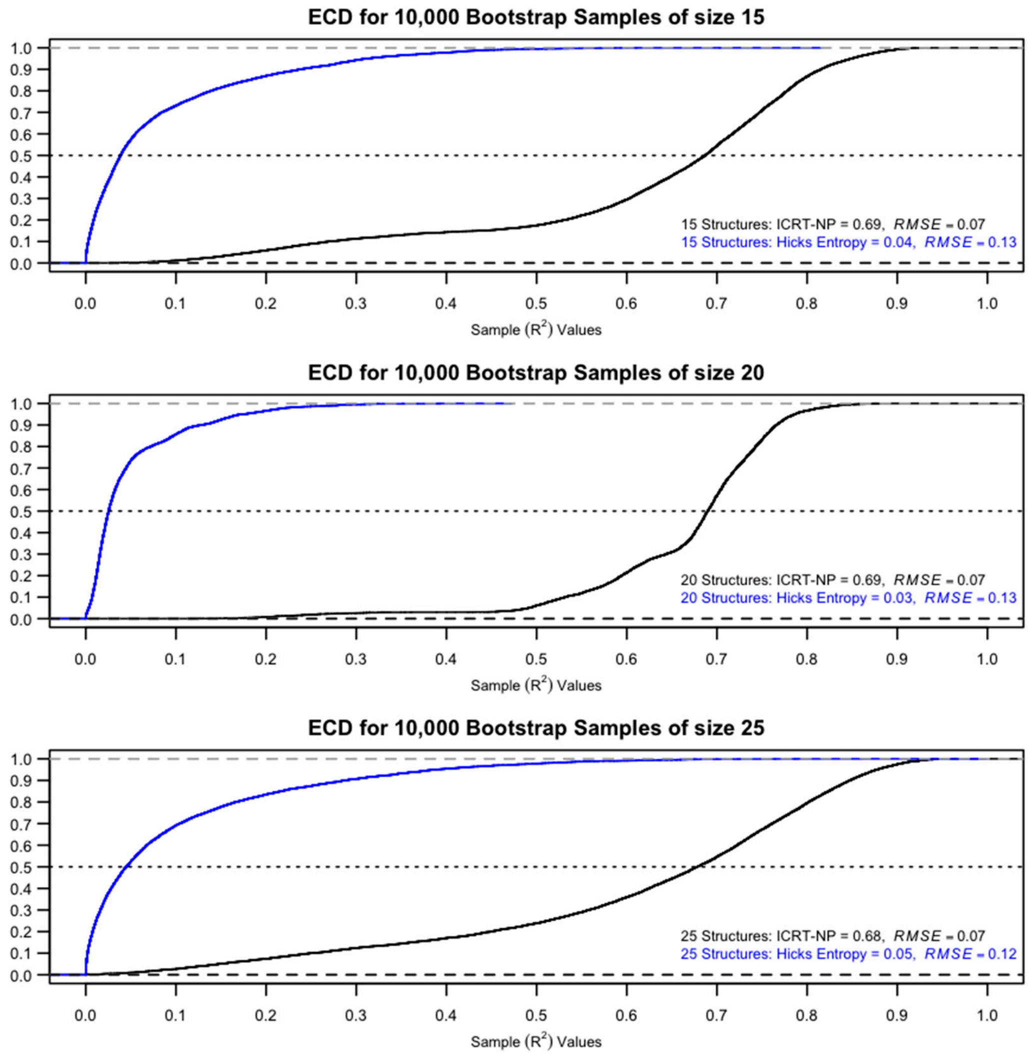

Figure 5. Bootstrapped linear regression analysis of the ICRT-NP and Hick's entropy-based law to account for choice reaction times associated with the four-object three-dimensional and four-dimensional category structures. The median R^2 and RMSE values are provided for each model within each plot, and the y-axis represents the cumulative proportion of R^2 values for each of the tested models. ECD = Empirical Cumulative Distribution.

9. Discussion

Hick's law was one of many attempts during the early development of Cognitive Science to apply Shannon information theory and its ubiquitous entropy measure to cognitive phenomena. Perhaps the most influential of these attempts may be found in the work on short-term memory by Miller [38] and the work on choice reaction times by Laming [5,6]. Yet, Laming later admitted the inadequacy of this approach with respect to his own work [6], while Miller's approach was found equally lacking [39,40]. Remarkably, despite these failures, even today's researchers continue to resort to Shannon entropy to account for complex high-level cognitive capacities. For example, most recently, Pape et al. [19] intrepidly attempted to account for concept learning difficulty orders using Shannon entropy. However, Vigo et al. [18] revealed multiple objective flaws with their approach.

What these failures suggest is the need for a new type of information theory that is suitable for psychological research. Undoubtedly, information is a powerful construct in its generality and relevance to cognitive research. However, reliance on a probabilistic and statistical notion of information that was originally intended for the analysis of communication between electronic devices does not serve the goal of describing the organic character of information and communication with respect to human observers. This forced approach conforms to the proverbial fallacy of fitting the square peg into the round hole. For this reason, Vigo [15,17,26] developed a new type of information theory better suited to the needs of cognitive research where context, as determined by the relationships between the object stimuli of categories, is of paramount importance. Similar to the ICRT, the theory—known as Representational Information Theory (RIT, and its generalized form as Generalized Representational Information Theory or GRIT)—is derived from the law of invariance in Generalized Invariance Structure Theory. Additionally, although there has been some research linking representational information to reaction times, it has been with respect to saccadic data as it pertains to classification behavior. Thus, the determination of a general connection between representational information and choice reaction times on structured sets warrants further research.

Instead, with the ICRT as an alternative to Hick's law, we have established a direct and natural link between the degree of concept learning difficulty of structured sets (i.e., categories) and choice reaction times that is supported by empirical results and that is not dependent on probability theory. This is important for several reasons: First, such a connection with respect to structured sets had never been established before, let alone non-probabilistically. Indeed, the structure-insensitive assumptions underlying probability theory and statistical modeling that have been inherited by Shannon information, as observed by Luce [4], are largely responsible for the failures of Shannon information with respect to psychological research. Accordingly, the key core construct in the ICRT law (i.e., degree of invariance) may be regarded as the non-probabilistic and structure-sensitive counterpart to Shannon entropy. This is because the degree of invariance of a category indicates its degree of orderliness from the standpoint of the relationships between its components (in other words, its logical coherence). Thus, zero degree of invariance indicates complete disorder on its corresponding deterministic scale.

Secondly, the underlying fundamental constructs of invariance, time, and category are unified under the banner of conceptual behavior—again, a first. Third, the scope and breadth of such integration speaks to its generality. Indeed, any situation that is analyzable in terms of categories (or categorization) may be linked to the choice reaction times on their elements. Fourth, the concept learning cognitive mechanism explaining the connection between choice reaction times and category structures has been precisely specified in previous work and provides a low-level causal link to decision-making phenomena. Fifth, GIST, the theory on which the ICRT and ICRT-NP are based, has been empirically tested on multiple occasions using historical data from well-known researchers (see [16,18,25,30]) and has been used by several researchers to inform their own work (see [41–43]).

Due to the generality of the ICRT-NP and the ICRT, their applications are far reaching. In the realm of cognitive psychology and decision sciences, predicting choice reaction times holds paramount importance—as do the possible mechanisms that explain choice reaction times. It is assumed that such predictions, if successful, may provide a greater understanding of the workings of the human mind, human brain, and everyday decision-making behavior. This greater understanding may, in turn, reveal ways of predicting human decisions with respect to different domains, including problem-solving. Accordingly, predicting reaction times may increase our understanding of the limits of human decision-making as a baseline for developing systems capable of emulating human intelligence.

Admittedly, a great deal of machine learning has focused on systems that can find optimal solutions without much concern for the types of limits on computational resources exhibited by the human brain (see [44,45]). However, an understanding of the cognitive processing limits associated with choice reaction times may facilitate the creation and characterization of *resourceful intelligent agents* that may be capable of processing small datasets more efficiently and in more humanly meaningful ways than any existing machine learning algorithm (for a discussion, see [18,26,45]).

10. Conclusions and Future Directions

In this paper, we have shown that the ICRT law and the ICRT-NP law account for choice reaction time data from Vigo and Doan [32] more accurately than both variants of Hick's law. The findings from both Figures 3 and 4 show the influence of relational information processing, independent of value-judgments or idiosyncratic preferences, on choice behavior. In other words, the fits using the ICRT and ICRT-NP laws suggest that the degree of categorical invariance that exists between category members captures the nature of the relational information processing underlying juxtaposed complex stimuli as is the case with these structured sets. These increases in degrees of categorical invariance between categories are reflected in Figures 3 and 4 by the increasing ICRT and ICRT-NP predictions on the x-axes.

Ecologically, this influence of relational information processing on choice behavior is ubiquitous in our daily lives in the sense that it is likely influential for any situation where one must compare a set of alternatives across a number of related dimensions or features. For example, when deciding which bag of coffee (or, more generally, any item) to purchase at the local supermarket, one must first compare each alternative across a finite number of relevant dimensions (e.g., brand, cost, roast type, size, etc.). Certainly, prior decisions and established preferences can influence decisions in such contexts. However, the invariance-based relational information processing that we suggest influences decision-making in our daily lives helps to establish these initial preferences, and then modifies them as new alternatives and/or dimensions are introduced into the decision-making situation.

Despite the ecological significance, there are some potential limitations with respect to the ICRT laws that should be noted. To start with, the laws have not been tested for structured sets or categories of objects defined over continuous dimensions. Second, the laws, in their current form, are probably most effective when, for any category X, $|X| > 2$ or $D > 2$. This is because the influence of meaningful interrelatedness or logical coherence between the features (and hence dimensions) of the object stimuli in a category—as measured by categorical invariance—increases as the number of object stimuli and/or the number of dimensions increase. This is the case at least until the number of objects or dimensions becomes intractable to process for a human observer. Thus, beyond just the number of objects (as Hick's law suggests) and/or the number of stimulus dimensions, the automatic and inevitable perceived key interactions between the objects of the category (as discovered by the proposed invariance pattern detection mechanism) also play a significant role in determining how long it takes to make a selection. Likewise, entropy-based models contemporary to Hick's and based on principles of probability theory are often ill-equipped to handle these perceived inter-object relations and will struggle to capture the conceptual character of choice reaction time.

Notably, these interactions are often too simple to influence the selection processing speed in a significant way, as in the case of choosing between only two alternatives with only two relevant dimensions. At this simplest level, an additional level of processing may occur that is purely based on subjective preferences and that involves the dimensions of quality, cost, utility, value, and so on that are typically proposed in applied decision-making theories. Building a bridge between such multiple levels of analysis is our current goal.

Finally, and importantly, both versions of the ICRT used in this paper were greatly handicapped in that they do not include potential free parameters to be found in the cognitive mechanism used to compute degree of invariance (see the definition of the lambda operator in the Supplementary Technical Materials). These potential free parameters–the *tau* discrimination threshold(s) and the invariance sensitivity/bias weights–were not used to account for group-level performance but may be used to meaningfully account for individual differences in choice reaction times between people. Given the basic goals of this paper, we did not include an analysis of individual differences. In a planned follow-up paper, we aim to demonstrate the effectiveness of the ICRT in accounting for individual differences. Indeed, based on the fairly accurate group level predictions achieved by the non-parametric ICRT, we expect that its free-parameter enriched counterpart will yield at least similarly accurate individual level predictions.

In conclusion, we demonstrated how reaction times from structured choice sets are well accounted for by an invariance-based law of conceptual behavior. In addition, we showed how Hick's law and the entropy-based application of Hick's law provide less accurate fits to the reaction time data. Although the evidence considered was only from one study, the study did involve the systematic and rigorous testing of over fifty logically distinct category structures across different groups of participants and assessed hundreds of reaction times per each of the category structures. Ultimately, we look forward to further testing the effectiveness and generality of the ICRT laws and continuing to build a bridge between conceptual behavior and decision-making.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/math11112422/s1. References [46–49] are cited in the supplementary materials.

Author Contributions: Conceptualization, R.V.; Methodology, R.V., C.A.D., J.W. and C.B.R.; Validation, C.A.D., J.W. and C.B.R.; Formal Analysis, R.V.; Investigation, R.V. and C.A.D.; Data curation, C.A.D. and J.W.; Writing—original draft, R.V.; Writing—review & editing, C.A.D., J.W. and C.B.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The aggregated reaction time data from Vigo and Doan [32] and the model predictions associated with Figures 3 and 4 can be accessed via the OSF website (https://osf.io/6eryv/) or the DOI: 10.17605/OSF.IO/6ERYV. Also available at this OSF website/DOI is the R code for performing the bootstrapping analysis reported in the manuscript and for generating the ECD plots in Figure 5.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hick, W.E. On the rate of gain of information. *Q. J. Exp. Psychol.* **1952**, *4*, 11–26. [CrossRef]
2. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [CrossRef]
3. Shannon, C.E.; Weaver, W. *The Mathematical Theory of Communication*; University of Illinois Press: Champaign, IL, USA, 1949.
4. Luce, R.D. Whatever happened to information theory in psychology? *Rev. Gen. Psychol.* **2003**, *7*, 183–188. [CrossRef]
5. Laming, D.R.J. *Information Theory of Choice-Reaction Times*; Academic Press: Cambridge, MA, USA, 1968.
6. Laming, D. Statistical information, uncertainty, and Bayes' theorem: Some applications in experimental psychology. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty, Proceedings of the 6th European Conference 2001, ECSQARU 2001, Toulouse, France, 19–21 September 2001*; Springer: Berlin/Heidelberg, Germany, 2001; pp. 635–646.
7. Jensen, A.R. *Clocking the Mind: Mental Chronometry and Individual Differences*; Elsevier: Oxford, UK, 2006.
8. Proctor, R.W.; Van Zandt, T. *Human Factors in Simple and Complex Systems*; CRC Press: Boca Raton, FL, USA, 2008.
9. Wickens, C.D.; Hollands, J.G.; Parasuraman, R.; Banbury, S. *Engineering Psychologyhuman Performance*; Routledge: Oxfordshire, UK, 2013.

10. Ali, A.; Liem, A. The use of formal aesthetic principles as a tool for design conceptualisation and detailing. In Proceedings of the NordDesign 2014, Espoo, Finland, 27–29 August 2014.
11. Chan, H.C.; Goswami, S.; Kim, H.W. An alternative fit through problem representation in cognitive fit theory. *J. Database Manag.* **2012**, *23*, 22–43. [CrossRef]
12. MacKenzie, I.S. *Human-Computer Interaction: An Empirical Perspective*; Morgan Kaufmann: Burlington, MA, USA, 2013.
13. Liu, W.; Gori, J.; Rioul, O.; Beaudouin-Lafon, M.; Guiard, Y. How relevant is Hick's law for HCI? In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020.
14. Proctor, R.W.; Schneider, D.W. Hick's law for choice reaction time: A review. *Q. J. Exp. Psychol.* **2018**, *71*, 1281–1299. [CrossRef]
15. Vigo, R. Complexity over uncertainty in generalized representational information theory (GRIT): A structure-sensitive general theory of information. *Information* **2012**, *4*, 1–30. [CrossRef]
16. Vigo, R. The GIST of concepts. *Cognition* **2013**, *129*, 138–162. [CrossRef] [PubMed]
17. Vigo, R. *Mathematical Principles of Human Conceptual Behavior: The Structural Nature of Conceptual Representation and Processing (Paperback ed.)*; Routledge: Oxfordshire, UK, 2015.
18. Vigo, R.; Doan, C.A.; Zhao, L. Classification of three-dimensional integral stimuli: Accounting for a replication and extension of Nosofsky and Palmeri (1996) with a dual discrimination invariance model. *J. Exp. Psychol. Learn. Mem. Cogn.* **2022**, *48*, 1165–1192. [CrossRef]
19. Pape, A.D.; Kurtz, K.J.; Sayama, H. Complexity measures and concept learning. *J. Math. Psychol.* **2015**, *64*, 66–75. [CrossRef]
20. Vigo, R. A Dialogue On Concepts. *Think* **2010**, *9*, 109–120. [CrossRef]
21. Kveraga, K.; Boucher, L.; Hughes, H.C. Saccades operate in violation of Hick's law. *Exp. Brain Res.* **2002**, *146*, 307–314. [CrossRef] [PubMed]
22. Pavão, R.; Savietto, J.P.; Sato, J.R.; Xavier, G.F.; Helene, A.F. On sequence learning models: Open-loop control not strictly guided by Hick's law. *Sci. Rep.* **2016**, *6*, 23018. [CrossRef] [PubMed]
23. Lawrence, B.M.; John, A.S.; Abrams, R.A.; Snyder, L.H. An anti-Hick's effect in monkey and human saccade reaction times. *J. Vis.* **2008**, *8*, 26. [CrossRef] [PubMed]
24. Wright, C.E.; Marino, V.F.; Belovsky, S.A.; Chubb, C. Visually guided, aimed movements can be unaffected by stimulus-response uncertainty. *Exp. Brain Res.* **2007**, *179*, 275–496. [CrossRef] [PubMed]
25. Vigo, R. Categorical invariance and structural complexity in human concept learning. *J. Math. Psychol.* **2009**, *53*, 203–221. [CrossRef]
26. Vigo, R. Representational information: A new general notion and measure of information. *Inf. Sci.* **2011**, *181*, 4847–4859. [CrossRef]
27. Vigo, R. Generalized Invariance Structure Theory: A Tutorial and Extensions, including a formal proof of the Similarity-Invariance Equivalence Principle. 2023; *in preparation*.
28. Shepard, R.N.; Hovland, C.I.; Jenkins, H.M. Learning and memorization of classifications. *Psychol. Monogr. Gen. Appl.* **1961**, *75*, 1. [CrossRef]
29. Nosofsky, R.M.; Palmeri, T.J. Learning to classify integral-dimension stimuli. *Psychon. Bull.* **1996**, *3*, 222–226. [CrossRef]
30. Vigo, R.; Evans, S.W.; Owens, J.S. Categorization behaviour in adults, adolescents, and attention-deficit/hyperactivity disorder adolescents: A comparative investigation. *Q. J. Exp. Psychol.* **2015**, *68*, 1058–1072. [CrossRef]
31. Vigo, R.; Doan, C.A.; Basawaraj; Zeigler, D.E. Context, structure, and informativeness judgments: An extensive empirical investigation. *Mem. Cogn.* **2020**, *48*, 1089–1111. [CrossRef]
32. Vigo, R.; Doan, C.A. The structure of choice. *Cogn. Syst. Res.* **2015**, *36*, 1–14. [CrossRef]
33. Goodwin, G.P.; Johnson-Laird, P.N. Mental models of Boolean concepts. *Cogn. Psychol.* **2011**, *63*, 34–59. [CrossRef] [PubMed]
34. Goodwin, G.P.; Johnson-Laird, P.N. The acquisition of Boolean concepts. *Trends Cogn. Sci.* **2013**, *17*, 128–133. [CrossRef]
35. Nosofsky, R.M.; Gluck, M.A.; Palmeri, T.J.; McKinley, S.C.; Glauthier, P. Comparing models of rule-based classification learning: A replication and extension of Shepard, Hovland, and Jenkins (1961). *Mem. Cogn.* **1994**, *22*, 352–369. [CrossRef] [PubMed]
36. Rehder, B.; Hoffman, A.B. Eyetracking and selective attention in category learning. *Cogn. Psychol.* **2005**, *51*, 1–41. [CrossRef] [PubMed]
37. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2020; Available online: http://www.R-project.org/ (accessed on 10 October 2020).
38. Miller, G.A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.* **1956**, *63*, 81. [CrossRef]
39. Shiffrin, R.M.; Nosofsky, R.M. Seven plus or minus two: A commentary on capacity limitations. *Psychol. Rev.* **1994**, *101*, 357–361. [CrossRef]
40. Cowan, N. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behav. Brain Sci.* **2001**, *24*, 87–114. [CrossRef]
41. Cai, X.; Li, F.; Wang, J.; Li, H. Invariance detection in the brain: Revealed in a stepwise category induction task. *Brain Res.* **2014**, *1575*, 55–65. [CrossRef]
42. Gao, H.; Cai, X.; Li, F.; Zhang, S.; Li, H. How the brain detects invariance and inhibits variance during category induction. *Neurosci. Lett.* **2016**, *626*, 174–181. [CrossRef]

43. Stoewer, P.; Schilling, A.; Maier, A.; Krauss, P. Neural network based formation of cognitive maps of semantic spaces and the putative emergence of abstract concepts. *Sci. Rep.* **2023**, *13*, 3644. [CrossRef] [PubMed]
44. Newell, A.; Simon, H.A. *Human Problem Solving*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1972; Volume 104.
45. Vigo, R.; Zeigler, D.E.; Wimsatt, J. Uncharted Aspects of Human Intelligence in Knowledge-Based "Intelligent" Systems. *Philosophies* **2022**, *7*, 46. [CrossRef]
46. Kruskal, J.B.; Wish, M. *Multidimensional Scaling*; Sage: Newcastle, UK, 1978; Volume 11.
47. Nosofsky, R.M. Choice, similarity, and the context theory of classification. *J. Exp. Psychol. Learn. Mem. Cogn.* **1984**, *10*, 104. [CrossRef] [PubMed]
48. Shepard, R.N. Toward a universal law of generalization for psychological science. *Science* **1987**, *237*, 1317–1323. [CrossRef] [PubMed]
49. Shepard, R.N.; Romney, A.K.; Nerlove, S.B. *Multidimensional Scaling: Theory and Applications in the Behavioral Sciences: I. Theory*; Seminar Press: Princeton, NJ, USA, 1972.