Data-Based Predictions: Holistic and Atomistic Procedures *

Jacob Glazer

Coller Faculty of Management, Tel Aviv University
and Department of Economics, Warwick University

Ariel Rubinstein

School of Economics, Tel Aviv University
and Department of Economics, New York University

July 4, 2021

ABSTRACT: Individual A would like to predict the choice individual B will make based on data about B's past choices in identical situations. We conduct a series of experiments in which subjects play the role of individual A and are asked to explain their prediction. We identify a variety of procedures, some of which relate to the prediction as a whole (*holistic*), while others relate to the various aspects of the prediction separately (*atomistic*). The subjects' predictions were found to depend on the way in which data is presented and to often differ from what one would expect based on conventional models.

AEA Classification: D01.

KEYWORDS: Experimental economics, non-Bayesian beliefs, Data-based predictions.

*We are grateful to Aron Tobias for a thorough review of the paper. We also thank Ayala Arad, Michael Crystal, Kfir Eliaz, Ignacio Esponda, Gerd Gigerenzer, Ori Heffetz, Ayush Pant, Dotan Persitz, Kirill Pogorelskiy, Andrew Schotter, Emanuel Vespa and Xueying Zhao for valuable comments.

1. Introduction

Individual A wishes to correctly predict the choice individual B will make. A lacks any knowledge of B's motives. He observes some data about B's past choices in the same choice problem. A succeeds if and only if his prediction is correct and fails if it is wrong.

The conventional way to model such a situation is to assume that A has in mind a probability belief over B's set of alternatives and chooses the one with the greatest likelihood. When, in addition, A also has partial information about the choice B is about to make, we conventionally assume that A predicts the alternative that is most likely, conditional on his partial information. However, in practice, A often does not have an explicit probability measure and can only base his prediction on the data he possesses about B's past choices in a number of past events.

Our purpose is to experimentally achieve some insight into procedures used by individuals to make a point prediction about another individual's behavior, based on some information about his past behavior in identical situations. The study consists of four sets of experiments. The experiments in each set begin with the same short description of a daily choice problem faced by a particular individual, followed by information on his past choices in identical situations. The subjects were then asked to provide their "best guess" as to the individual's choice on the following day by selecting one possibility from an alphabetically ordered list of all possibilities. After submitting their prediction, subjects were simply asked to "please explain your choice". A majority of the subjects provided fairly comprehensive explanations which enable us to classify them according to a typology of data-based prediction procedures.

We confine ourselves to scenarios in which individual A can either succeed or fail in achieving his goal, namely to predict the most likely event. Restricting ourselves to win/loss situations avoids becoming entangled in theories of decision making under uncertainty.

The experiments were conducted online by means of the site gametheory.tau.ac.il[1] which is used mainly for pedagogical purposes in game theory courses around the world. The subjects were all students who had taken or were taking a game theory course. In each experiment, a subject only had to provide his e-mail address, make one prediction and explain that prediction on a subsequent screen. No control questions were asked and the subject were not asked to provide any additional information. Some of the students participated in more than one set of experiments but not more than one experiment in each set. In each set, four subjects who had provided a prediction and an explanation were drawn randomly and independently of their answers were awarded $40. No other incentives were provided. The subjects – the vast majority of whom explained their choices – appear to have taken the experiment quite seriously, as can be

[1] Currently, https://arielrubinstein.org/gt/ .

seen from the detailed explanations they provided. Notice that incentivizing the subjects here was problematic because in none of our experiments there is a "correct" prediction and there was no real randomizing process that generated the data.

In each experiment, all the alternatives have the same structure combining a number of components that are to be chosen simultaneously. We distinguish between *holistic* and *atomistic* procedures, both of which were observed in all the experiments:
Holistic – In making his prediction, a subject views each alternative as a whole.
Atomistic – In making his prediction, a subject deliberates on each component separately.

To illustrate, consider the following question, presented to (half of) the subjects in the set of experiments discussed in the next section:
"A flower store sells 4 types of flowers: G, O, P and R. Tom is a regular customer. Every Friday he comes to the store and buys a bouquet consisting of two types of flowers. The following table specifies how many times he bought each pair of flowers during the last 62 weeks:

pair	GO	GP	GR	OP	OR	PR
# of times	11	15	7	8	3	18

What is your best guess as to the pair of flowers he will choose next Friday?
In this experiment, predicting the most frequently chosen pair of flowers (PR) is a prime example of a holistic procedure (later to be referred to as H^{max}). An example of an atomistic procedure (to be referred to as T^{max}) would be to choose the bouquet consisting of the two most frequently chosen flowers (GP).

In all of the experiments, the subjects had the possibility of deriving the distribution of past choices. Nevertheless, it was found that often they did not choose the most frequent alternative. This can be attributed to three factors:

(i) Even when the data specified the number of times each alternative was chosen in the past, many subjects used atomistic procedures which do not necessarily lead to the most frequently observed alternative.

(ii) When the data was presented in such a way that the subjects had to derive the distribution of past choices, many used atomistic procedures or followed rules of thumb that do not necessarily lead to the most frequent past choice (for a survey of decision-making heuristics, see Gigerenzer and Gaissmaier (2011)). This is not a standard framing effect a la Tversky and Kahneman (1986) (see the discussion in Section 7).

(iii) Even when the subjects were aware of the distribution of past choices, a small but significant number did not choose the most frequent one, and in some cases they even chose the least frequent one. These subjects believed that people tend to diversify their choices or that their behavior is governed by a random variable and employed a "gambler's fallacy"-type of consideration in making their predictions (Kahneman and Tversky (1971)).

The core of the paper presents four sets of experiments, each designed to demonstrate a different point. We report on each set in isolation using the same format: after describing the set's experiments, we detail the more commonly used procedures of choice, each of which is classified as holistic or atomistic, and provide their observed frequencies. The observed distributions of responses should be treated only as suggestive of the procedures commonly used by people in such situations and are at most very rough approximations of the actual distributions of such procedures. The discussion of the four sets is followed by comments on the rationale of some of the observed procedures. The related literature is discussed in the final section. The purpose of this paper is not to come up with a unified theory of how people make data based predictions about other people's behavior. Our goal three fold: to demonstrate that even in very simple situations different individuals may apply very different procedures and come up with very different predications; to point out to some of the more common procedures that individuals use and to show how sensitive these procedures are to the type of data presented and to the way in which it is presented.

We are not attempting to arrive at a single model that can describe an individual's behavior when making a data-based prediction of another individual's behavior, a common task in the real world. The results demonstrate that even in a very simple situation individuals vary in the procedures they use. Nevertheless, the results point to some common procedural features in an individual's deliberation and the type of data that is available has a decisive impact on an individual's prediction. Given that the data is often provided by an interested party, the results may be used to explain which data is presented. In any case, the challenge is to integrate our observations into an economic model of interaction between agents, a task that is far from being straightforward and is left to future research.

2. A bouquet of flowers (DG 6+7)[2]

In this set of experiments, the data is presented to the subjects in the form of a table which provides the number of days on which each alternative was chosen. We refer to this type of data as a *distribution of choices*.

2.1 The experiments

The two variants of the experiment start as follows:
"A flower store sells 4 types of flowers: G, O, P and R. Tom is a regular customer. Every Friday he comes to the store and buys a bouquet consisting of two types of flowers. The table below specifies how many times he bought each pair of flowers during the last (44/62) weeks:"
The subjects were then randomly presented with one of two distributions of possible pairs of flower types. The two sets of data, denoted by B1 and B2, are presented in Table 1:

	B1	B2
GO	11	11
GP	15	15
GR	-	7
OP	-	8
OR	-	3
PR	18	18

Table 1: B-data

We concluded with the question: "What is your best guess as to the pair of flowers he will choose next Friday?"

[2]These symbols are our administrative notations.

2.2 Results

Based on the subjects' explanations of their predictions, we identify two main holistic procedures (the choice generated by each procedure appears in parentheses):
H^{max}: Choose the most frequently chosen combination (PR).
H^{min}: Choose an alternative different from the most frequently chosen one. This type of procedure has several versions, such as "choose the least frequently chosen alternative from among those that were chosen at least once" (GO in B1 and OR in B2) or "choose the second-most frequent alternative" (GP).

The two main atomistic procedures were: [3]
T^{max}: Choose the bouquet consisting of the two most frequently chosen colors (GP).
T^{min}: Choose the bouquet consisting of the two least frequently chosen colors (OR).

Table 2 presents (again) the data given to the subjects alongside the results. Each entry corresponds to a particular choice made in a particular version of the experiment. The results are reported after excluding subjects who either: (i) gave no explanation for their choice; (ii) gave an explanation for a choice other than the one they actually made; or (iii) gave a nonsensical explanation (such as "bla, bla, bla"). We indicate the number of subjects excluded from each entry by ⟨...⟩. The last row of the table shows the total number of omitted answers and their share of the *total* number of answers.

	B1	B1 (results)	B2	B2 (results)
GO	11	11% [18: 12H^{min}]⟨4⟩	11	4% [5: 5H^{min}] ⟨11⟩
GP	15	17% [27: 25T^{max}, 2H^{min}]⟨7⟩	15	32% [41: 32T^{max}, 5H^{min}] ⟨5⟩
GR	-	3% [5:]⟨3⟩	7	2% [2:]⟨3⟩
OP	-	1% [1:]⟨3⟩	8	1% [1:]⟨4⟩
OR	-	4% [6: 4T^{min}] ⟨2⟩	3	8% [10: 10H^{min}]⟨2⟩
PR	18	64% [102: 98H^{max}] ⟨7⟩	18	54% [70: 67H^{max}] ⟨5⟩
n=		[159]⟨26=14%⟩		[129]⟨30=19%⟩

Table 2: B - Results

In each entry, a larger font indicates the percentage of subjects participating in the experiment that corresponds to the entry's column and who gave the answer that corresponds to the entry's row. The first element in the square brackets is the number of subjects who gave the answer corresponding to that entry and it is followed by statistics about the main procedures used to explain the corresponding choice. For example, the entry at row GP and column B2 indicates that 32% of the subjects in B2 gave the answer GP; this group consisted of 41 subjects, of whom 32 used T^{max} while 5 used H^{min}. As

[3] Very few subjects followed some other atomistic procedure, such as: first choose the most frequently chosen color and then pair it with its most frequently chosen partner, which leads to PR.

indicated in the diamond parentheses, 5 answers of GP were omitted in B2. The entry at row GR and column B1 indicates that 5 subjects chose GR in B1 and that none of them gave an explanation that fits into any of our main categories of procedures.

In both versions, a majority of subjects applied H^{max} (PR). A significant share of the subjects chose GP (the two most frequent colors) and explained their choice using T^{max}. In B1, 62% of the subjects used H^{max} while only 15% used T^{max}. In B2, there was a significant decline to 52% in the use of H^{max} and an increase to 25% in the use of T^{max}. The difference is likely due to the fact that B1 induces the use of a holistic procedure since only three of the six combinations are present in the distribution of past choices, while in B2 it is less evident that the individual's choices were made holistically. About 10% of the subjects mentioned minimization rather than maximization in their explanations. Most of them chose the least frequently observed combination from those that were observed a non-negligible number of times. The fact that we observe H^{min} much more frequently than T^{min} may be related to Simonson (1990)'s claim that "people tend to choose more diversity when the choices are bracketed broadly than when they are bracketed narrowly".

Response Time: We also measured the subjects' response time (the time between when they received the question and when they submitted their choice). Response time is often useful in interpreting the choice made and, in particular, distinguishing between instinctive and contemplative responses (see, for example, Rubinstein (2007, 2013)). Given the relatively small number of subjects in the experiments, deriving conclusions from response time is of limited value. Nevertheless, it is interesting to note that in B1, there is no significant difference in median response time (MRT) between the subjects who chose PR (63s) and those who chose GP (70s) or GO (67s). In contrast, the MRT of the subjects in B2 who chose GP (103s) is much longer than for those who chose PR (65s). This is likely due to the fact that the execution of the atomistic T^{max} procedure, which leads to GP, requires more effort in B2 than in B1.

3. Web path (DG 4)

The next set consists of four experiments, all of which start with the following background story: "An individual starts his morning by surfing the web. He visits each of the three websites A, B and C exactly once. There are links from each site to the other two. The order in which he chooses to visit the three sites can vary. Recently, the individual was tracked for 19 (21) days." The story is followed by some data and then subjects were asked: "What is your best guess as to his route tomorrow?"

In experiments V1 and V1*, the data is presented as a distribution of choices, as it was in the B-experiments reported in Section 2. These experiments are necessary in order to provide a comparison to V2 and V3, in which the data is presented in a different form, although the underlying choice distribution remains the same as in V1 and V1*, respectively.

3.1 V1 and V1 – the experiments*

In V1 and V1*, the background story was followed by: "... and the route he used each day was recorded. The following table indicates the number of times he followed each of the six possible routes:". Table 3 presents the distributions of choices alongside the results.

The difference between these experiments and the previous ones lies in the structure of the alternatives: here an alternative is a permutation of 3 elements (A, B and C) whereas in the B-experiments it was a subset of two of the four elements G, O, P and R. A holistic procedure treats each permutation as a whole, while an atomistic procedure starts by guessing a particular website (usually the first) and then uses some rule to complete the predicted path.

As in the previous section the holistic procedures H^{max} and H^{min} predict the most observed and least observed route respectively. In this section, we classify as T^{max} all procedures that predict a path starting from the site that appears most frequently as the first site and predict the second site according to some consideration.

In V1, H^{max} leads to BAC. The H^{min} procedure in its stronger version leads to one of the two unchosen permutations, ACB or CBA[4], and in its weaker version to either a permutation that has not yet been selected, ACB or CBA, or to the least frequent permutation with a positive number of selections BCA.[5] The procedure T^{max} predicts B as the first site. A few subjects predict A as the second site, either because it is the most frequent site conditional on B being the first site[6] or because A is the most frequent second site.[7] Some subjects proceeded by predicting C as the second site based on a variety of explanations or without providing any explanation at all. The procedure T^{min} leads to the avoidance of B as the first site.

The conflict between holistic and atomistic procedures is more pronounced in V1* where H^{max} leads to CAB while T^{max} leads to the choice of B as the first site. After predicting B as the starting point, the T^{max} users split: some chose BCA since, conditional on B being the first site, C is the most likely second site[8] while others chose BAC since

[4] ACB: "I had to choose either from ACB or CBA, as they were the ones still not used."

[5] BCA: "The least frequent of the most frequent ones."

[6] BAC: "He visits B first on 10 out of 19 times. Moreover, when he visited B first, on 7 out of 10 times he visited A next then C. BAC looks like the most frequent option".

[7] BAC: "The individual does not seem to be indifferent between the choices. His choice B for the first website highly dominates. For his second choice the AC seems to be dominant again...".

[8] BCA: "This is a sequential event. He is more likely to pick B first. Given picking B, he is then more likely to pick C afterwards."

A is the most frequently chosen second site.[9] Explanations that refer only to B as being the first site are also classified as T^{max}.

3.2 V1 and V1 – results*

Following are the results for V1 and V1* (see Section 2.2 for an explanation of how to interpret the entries).

	V1	V1 (results)	V1*	V1* (results)
ABC	4	5%[4: 3H^{min}]⟨2⟩	0	0% [0:]⟨2⟩
ACB	0	4% [3: H^{min}]⟨0⟩	0	1% [1:]⟨2⟩
BAC	7	73% [56: 51H^{max}, 5T^{max}]⟨6⟩	6	27% [23: 12T^{max}, 6H^{min}]⟨2⟩
BCA	3	14% [11: 1H^{min}, 6T^{max}]⟨3⟩	7	30% [26: 22T^{max}]⟨4⟩
CAB	5	3% [2: 2H^{min}]⟨4⟩	8	42% [36: 36H^{max}]⟨6⟩
CBA	0	1% [1: 1H^{min}]⟨0⟩	0	0% [0:] ⟨0⟩
n=		[77] ⟨15=16%⟩		[86]⟨16=16%⟩

Table 3: V1 and V1* – results.

In V1, both H^{max} and one version of T^{max} lead to the choice of BAC. Unsurprisingly, the vast majority of subjects (73%) chose this alternative, with almost all of them using H^{max}. All together only 15% of the subjects used T^{max}, choosing either BAC or BCA.

Faced with the dilemma of choosing between holistic and atomistic procedures in V1*, fewer subjects used H^{max} than in V1 (a decline from 66% to 42%) while more of them used T^{max} (an increase from 14% to 40%). In both V1 and V1*, about 10% of the subjects used H^{min}.

3.3 Data presenttion – transitions

In V2, the data was presented in a more implicit manner than by means of a distribution of choices. The subjects were provided with the number of days on which the individual had used each of the 6 possible links (from one site to another). The corresponding distribution of choices is the same as in V1.

[9] BAC: "If I consider each letter in each position then B first is the most common and A second. Which leaves C with only one option as third."

Specifically, the data was presented in the following form: "Each day, the two links he used were recorded. The following table presents the number of times he used each of the six links whether as the first or the second:"

		to	
from	A	B	C
A	-	9	7
B	7	-	7
C	8	0	-

Table 4: V2-data

Note that the sum of the six numbers in the table must be 38 and the table always corresponds to a unique distribution of choices (in this case, that in V1).[10]

3.4 V2 (transitions) – results

We identified the following commonly used procedures:

Holistic:
H^{max}: Solve the underlying distribution of choices and choose the most frequent route (BAC). In this particular case, finding the distribution of choices is non-trivial though not particularly difficult either.[11]
H^{sum}: Choose the route XYZ with the highest number of transitions from X to Y and from Y to Z (CAB).
H^{likely}: Some subjects justified the choice of ABC as being "the most likely path".

Atomistic:
$T1$: Start with the pair XY which has the largest number attached to the link from X to Y (ABC).[12]
$T2$: Start with the bold observation that C never leads to B. Conclude (wrongly) that the sequence must include CA. Then, note that the transition from A to B is more likely than the transition from B to C and conclude the sequence CAB.[13]

[10] Let T be the set of the six links $(x, y) \in X \times X$ where $x \neq y$. Let P be the set of the six paths and let M be the matrix of size $|T| \times |P|$ where $M_{tp} = 1$ if the transition t is on the path P and 0 otherwise. Given the distribution of choices (s_p), the data (s_t) is obtained by the formula $(s_t) = M(s_p)$. Since the matrix M is invertible, the distribution of choices is fully revealed by the frequencies of the links.

[11] First, note that the individual must have started at A on 4 occasions. Then, it must be that he followed the path ABC on 4 occasions since there are no transitions from C to B .

[12] ABC : "from A to B he moves the most, from B to C is the second possible move."

[13] CAB: "From C to A is the most frequent choice involving C and then from A she goes to B. This step also happened more frequently than her going from B to C first."

$T3$: Contract the sequence striating with CA affected by the bold observation that C has never been followed by B.
$T4$: This category includes several justifications for choosing B first and C last.[14]

Note that $T1$ and $T3$ are holistic in the sense that the choice of a segment determines the entire path. We classify them as atomistic since they are holistic only in the case of 3-site paths.

Following are the results for V2 presented alongside to those of V1.

	V1	V2
ABC	5%	32%[29:9H^{likely}, 10$T1$] ⟨10⟩
ACB	4%	2% [2:] ⟨6⟩
BAC	73%	11%[10:4H^{max}, 4$T4$] ⟨7⟩
BCA	14%	3% [3:] ⟨3⟩
CAB	3%	51% [47: 33H^{sum}, 7$T2$, 4$T3$] ⟨8⟩
CBA	1%	1% [1:] ⟨2⟩
$n =$	[77]	[92] ⟨36=28%⟩

Table 5: V2-results

Note the striking difference in behavior between V1 and V2 even though the distribution of choices in V1 corresponds to that underlying the data in V2. A vast majority of the subjects chose BAC in V1 while a majority chose CAB in V2, most of them by using the H^{sum} procedure. H^{sum} is often (though not in this case) a rule of thumb that leads to the most frequent path (see Section 6). About 1/3 of the subjects used various types of atomistic procedures. Only one justified his choice by an argument involving minimization.

3.5 V3 (order of visits)

In V3, we provided the number of times each site was visited first, second and third. In this case, the corresponding distribution of choices was as in V1*. The data was provided as follows:

[14] BAC: "B has few incoming and many outgoing edges, therefore it is likely to be the starting site. C has few outgoing and many incoming edges, therefore it is likely to be the final site."

"Recently, the individual was tracked for 21 days. Each row in the following table indicates the number of days on which he visited the site first, second and third."

	1st	2nd	3rd
A	0	14	7
B	13	0	8
C	8	7	6

Table 6: V3 - data

Such a table has 4 degrees of freedom (since each row and each column must add up to the total number of observations) and it often corresponds to more than one distribution of choices. However, this particular set of data is consistent only with the distribution of choices in V1*.

3.6 Results – V3 (order of visits)

We identified the following main procedures:
Holistic:
H^{max}: Solve for the distribution of choices and choose the most frequent path (CAB).
H^{sum}: Choose the path (a_1, a_2, a_3) with the largest sum of numbers in the three entries $(a_1, 1)$, $(a_2, 2)$ and $(a_3, 3)$ (BAC). (Some subjects multiplied the entries rather than summing them.)
Atomistic:
T^{max}: Start with the site visited first most often. Proceed to the site with the highest number of second visits from among the remaining two sites (BAC).

	V1*	V3
ABC	0	0[0:] ⟨5⟩
ACB	1%	2%[2:] ⟨4⟩
BAC	27%	81%[62: 46T^{max}, 4H^{sum}] ⟨20⟩
BCA	30%	2%[2:] ⟨5⟩
CAB	42%	14% [11: 7 H^{max}] ⟨3⟩
CBA	0	0[0:] ⟨1⟩
n=	[86]	[77] ⟨38=33%⟩

Table 7: V3 - results

Although the data in V3 is only consistent with the distribution of choices in V1*, the results are strikingly different: 81% of the subjects chose BAC in V3 as opposed to only 27% in V1*.

The data in V3 was the most difficult for the subjects to process. About 1/3 of the subjects did not provide any explanation or they provided a nonsensical one. There was also an exceptionally large number of participants whose explanations could not be clearly classified. The dominant explanation was the atomistic T^{max}. Only 14% gave a clear holistic explanation: 7 subjects found the underlying distribution of choices and chose the most frequent sequence while 4 used H^{sum}.

3.7 Comments

Only one subject in V2 and two in V3 used some type of minimization procedure, which is in contrast to V1 and V1* where about 10% applied H^{min}. This raises the possibility that when the data is presented implicitly, subjects are "too busy" processing it to use a less instinctive procedure which avoids the most frequent alternative.

Note that there is a significant difference between the procedures used in V2 and V3. The most popular procedure in V2 is H^{sum} while in V3 it is T^{max}. However, we do not have any evidence that the type of data (V2 vs. V3) has a significant effect on whether subjects use holistic or atomistic procedures and the difference may be a result of the particular sets of data used in these experiments.

Response Time: There is insufficient data in order to analyse the response time in V1: Regarding V1*, the differences in MRT between the three major responses are insignificant (BAC (78s), BCA (82s) and CAB (71s)). The MRTs in V2 are: BAC (178s), CAB (156s) and ABC (109s) and in V3 they are: CAB (154s) and BAC (110s). Unsurprisingly, the MRTs in V2 and V3 that were the result of using H^{max} (finding the most frequent path in the underlying distribution of choices) were much higher than the rest.

4. Preferences (DG5)

This set of experiments is similar to the V experiments in that each alternative is still a sequence of three elements, but whereas in V the alternative xyz was a web path, here it is interpreted as the preference relation $x \succ y \succ z$. A data set describes the preferences of 20 individuals. W1 is analogous to V1 and consists of a distribution of rankings. W3 is analogous to V3 and presents the number of individuals in the group who rank each of the three elements as first, second or third. The exact analogue to V2 would report, for each pair of alternatives x and y, the number of individuals who rank x *just* above y.

However, this is not an intuitive formulation and therefore we replaced it in W2 with the number of individuals who prefer x to y for each pair of elements x and y.

All W versions start as follows: "Members of a community have interests in Art (A), Politics (P) and Sports (S) but attribute different priorities to each. Recently, 20 individuals from the community were surveyed and asked to rank the three interests according to level of priority." All three versions then present the data, each in a different form. They all conclude with: "Suppose that another individual is randomly selected from the community. What is your best guess as to his ranking of the priorities?"

4.1 W1 – distribution of rankings

The data set consists of the distribution of rankings presented in the column W1 in Table 10. W1 is similar to V1* in that it creates a conflict between holistic and atomistic procedures. In this case, the only frequently used holistic procedure is H^{max}, which leads to APS. We identified several atomistic procedures:
T^{12}: Rank as first the most frequent first-priority interest and rank as second the most frequent second-priority interest (PSA).[15]
T^{13}: Rank as first the most frequent first-priority interest and rank as last the most frequent last-priority interest (PAS).[16]
T^{P}: Rank the interests according to their frequencies as the first priority (PAS).[17]

4.2 W2 – binary comparisons

The data in W2 was presented as follows: "Each entry in the following table presents the number of individuals who assign a higher priority to the row interest over the column interest. (For example, 8 individuals assign a higher priority to A than to P.)"

	A	P	S
A		8	13
P	12		20
S	7	0	

Table 8: W2 - data

The data for the binary comparison of three alternatives is often consistent with

[15] PSA: "No one starts with S. 12/20 start with P. 7/12 choose S second."
[16] PAS: "P was the most liked one, while S was the most unliked one. A was in between."
[17] PAS: "There are more individuals that prioritize politics in this community than art (5+7>8), so I guess so does the next member. Art seems to be the second most important, so I chose the answer PAS."

more than one distribution of rankings.[18] However, the numbers in Table 8 are generated from the distribution of rankings given in W1 (presented in Table 10) and were chosen so as to be consistent with a unique distribution of rankings that can easily be derived from Table 8.

Only one holistic procedure was identified:

H^{max}: Find the underlying distribution of rankings and choose the most frequent ordering (APS).

Two notable atomistic procedures were identified:

$T^{\#wins}$: Rank the interests according to the number of "wins" over the other interests (PAS).

$T^{majority}$: Rank any pair of interests according to the majority rule (which in this case yields the transitive relation PAS) .

4.3 W3 (priorities)

The data in W3 was presented as follows: "The following table presents the number of individuals who rank each interest as being their first, second or third priority."

	1st priority	2nd priority	3rd priority
A	8	5	7
P	12	8	0
S	0	7	13

Table 9: W3 - data

The following holistic procedures were identified:

H^{max}: Find the distribution of rankings underlying the data and choose the most frequent ordering (APS).

H^{sum}: Choose the ordering (a_1, a_2, a_3) with the largest sum of the entries $(a_1, 1)$, $(a_2, 2)$ and $(a_3, 3)$ (PAS).

Two atomistic procedures were identified:

T^{12}: Rank as first the most popular first-priority interest. Rank as second the most popular second-priority interest from among the remaining two (PSA).

$T^{noticeable}$: Locate the two largest numbers in the table and if they appear in different columns then determine the order of interests accordingly (PAS).

[18] Let T be the set of six pairs $(x, y) \in X \times X$ where $x \neq y$. Let P be the set of six rankings and let M be the matrix of size $|T| \times |P|$ where $M_{tp} = 1$ if $t = (x, y)$ and x is preferred to y according to the ranking p and 0 otherwise. Given the distribution of rankings (s_p), the data (s_t) is obtained by the formula $(s_t) = M(s_p)$. The matrix M is of rank 4 and thus typically the distribution of rankings is not fully revealed by the frequencies of the binary comparisons.

4.4 W–results

DG5	W1	W1 - distribution	W2 - binary comparisons	W3 - priorities
APS	8	54% [55: 50H^{max}] ⟨6⟩	14% [10: 7H^{max}] ⟨4⟩	23%[19: 12H^{max}] ⟨8⟩
ASP	0	2% [2:] ⟨1⟩	0%[0:] ⟨1⟩	0% [0:] ⟨3⟩
PAS	5	20% [20: 10T^{P}, 5T^{13}] ⟨1⟩	74%[55: 38$T^{majority}$, 6$T^{\#wins}$] ⟨18⟩	57%[48: 2H^{sum}, 36 $T^{noticeable}$] ⟨8⟩
PSA	7	22% [22: 16T^{12}] ⟨2⟩	1%[1:] ⟨7⟩	17%[14: 12T^{12}] ⟨6⟩
SAP	0	1%[1:] ⟨2⟩	7%[5:] ⟨2⟩	1% [1:] ⟨2⟩
SPA	0	1% [1:] ⟨2⟩	4% [3:] ⟨5⟩	2% [2:] ⟨1⟩
n =		101 ⟨14=12%⟩	74 ⟨37=33%⟩	84 ⟨28=25%⟩

Table 10: W - results

The distribution of rankings in W1 is similar to that in V1*. About half of the subjects chose the most frequent ordering while the rest were split between the two orderings in which the most popular interest is the first priority.

In W2, we again observed that the presentation of the data is a critical factor in determining the predictions. Thus, in W2 significantly fewer subjects chose the most popular ranking, i.e. APS, while many more used an atomistic procedure which led them to PAS. Notice that the data in this case is such that the majority rule is transitive. We conjecture that if the data was such that the majority rule had a cycle, then more subjects would have used $T^{\#wins}$.

Although W1 is the only distribution of rankings consistent with W3, it leads to a drastically different distribution of predictions. In W3, as in V3, a small proportion of the subjects (14%) arrived at the distribution of rankings using H^{max} while very few applied the rule of thumb H^{sum}. Most subjects used one of the two atomistic procedures T^{12} or $T^{noticeable}$.

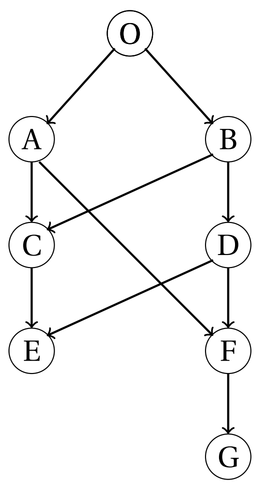
Comment: The information in the set W experiments for a sample of individuals. In contrast, the data in sets B and V refer to the past behavior of a single individual, which the subjects may believe that wishes to diversify his choices. This may explain the fact that in contrast to sets B and V there are hardly any subjects using H^{min} or T^{min} in set W.

Response Time: Given the relatively small number of subjects, conclusions based on response time have limited value. Nevertheless, it is interesting to note the following (intuitively appealing) observations: In W1, the instinctive prediction is APS (the most popular preference relation). Its median response time (MRT) was 74s whereas that of both PAS and PSA (the outcomes of the atomistic procedures) was 108s. In W2 and W3, the prediction of APS required a non-trivial calculation and the MRT of the small groups of subjects who predicted it was much longer than for those who predicted PAS, the outcome of the atomistic procedures (219s vs. 137s in W2 and 183s vs. 120s in W3).

5. The Graph Experiment (Z) (DG 2+3)

The background story of Z is similar to that of V: an individual chooses a path for surfing the web each day. However, in this experiment:
(a) the set of feasible paths is restricted to a given directed graph; and
(b) the subjects were provided with partial information about the individual's path on the day for which they were asked to make the prediction.

More precisely, the three versions of Z start with:
"An individual surfs the web every day. He always starts from O and finishes at either the website E or the website G. A link from one website to another is indicated in the graph by an arrow. Last month the individual was tracked for 29 days." There then appears the directed graph of Figure 1 which describes the feasibility constraints, together with data describing the individual's last 29 choices.



ACE	4
AFG	13
BCE	5
BDE	6
BDFG	1

Figure 1: Z1 – data

Finally, the subjects were asked the following: "Today it was observed that the individual ended up at E. What is your guess as to the individual's route today?" The three versions differ in the data presented to the subjects, but the distribution of choices was kept identical in all three.

We now present each of the three versions together with a description of the main procedures derived from the subjects' explanations. The distribution of answers and procedures is presented in the last subsection.

5.1 *Z1 – distribution of choices*

The basic version presents the distribution of choices as appears in Figure 1. We identified two holistic procedures: H^{max} (BDE) and H^{min} (ACE), and three atomistic procedures:

T^{max}: This procedure has two variants. In the first, select the most frequent first-visited site (A or B) unconditioned on the information that the path has to end at E and then proceed according to some rule (in this case, the procedure necessarily leads to the choice of ACE). In the second variant, select the most frequent first-visited site conditional on ending at E and then proceed according to some rule (which leads to either BCE or BDE).

T^{rep} : For each of the two possible first-visited sites, identify the "representative" terminal site. Start from a site for which E is a representative terminal site.[19]

T^{bi}: Start with E and move backwards to the most likely site (BCE).

5.2 *Z2 – transitions*

Z2 is parallel to V2. The subjects were provided with the number of transitions from one site to another: "The number next to an arrow is the number of times the person used the corresponding link."

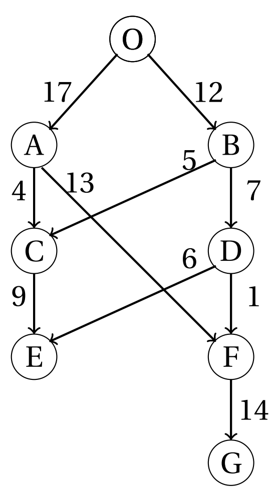


Figure 2: Z2 - data

[19] BDE: "When the individual chooses A web, then most of the times, he also chooses F web (which leads to G web). When the individual chooses B web, then most of the times, he also chooses C web (which leads to E web)".

We identified two *holistic* procedures:
H^{max}: Find the distribution of choices and choose the most frequent path (BDE).
H^{sum}: Choose the path ending at E with the maximal number of transitions (ACE).[20]

We identified the following *atomistic* procedures:
T^{max}: Proceed from the origin by choosing the more frequent transition, as long as that does not rule out ending at E (ACE). [21]
T^{rep}: Choose the first node which, according to the data, leads to E most frequently (BCE or BDE). [22]
T^{bi}: Move backwards from E by choosing the most frequent transition each time (BCE).[23] We include in this category any explanation that involves starting from E and moving to C even if choosing the rest of the sites involved other procedures.

5.3 Z3 – visits

In the third variant within the Z-set, subjects were provided with the number of times the individual visited each of the sites, along with the following text and chart: "The number next to each website is the number of times the individual visited this website on his 29 daily routes."

[20] ACE: "If you sum the amount of times he used each website in every row OACE-30, OBCE-26, OBDE-25, so i thought it was more likely that the rote would be OACE."

[21] ACE: "He chose A more days so the chance he will do it again is higher and from A it is the only way."

[22] BDE: "Out of 12 times he reaches to B he got 11 times to E, and just 4 times out of 17 he got from A to E. Then 7 times to D (instead of 5 times to C) and then E. better odds."

[23] BCE : "Out of the 15 times ending up at E, 9 times he got there by C and 6 times by D. Out of 9 times ending up at C, 5 times he got there by B, 4 times by 4. Therefore the most probable route was that from O he went to B, then C and then E."

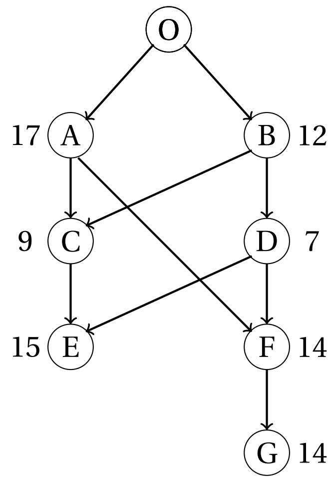


Figure 3: Z3 – data

In Z3, we observed procedures analogous to those in Z2: holistic – H^{max} (BDE) and H^{sum} (ACE)[24]; atomistic – T^{max} (ACE)[25], T^{rep} (BCE)[26] and T^{bi} (ACE).[27] For example, H^{sum} became: Choose the path ending at E with the largest sum of visits.

[24] ACE: "If I add up the amount of times the surfer visited in each site for each route, the route I chose has the biggest number of visits so it's most likely."

[25]ACE: "I picked the one he is more likely to pick statistically, (17/29 >12/29), and then the road was known (the only way to get to E from A is ACE)."

[26]BCE : "If he goes to A it is more likely he'll go to F. And because we know he went to E I guess he went to B first. It seems more likely he will end up to E from B. Because from C and D he can go to E still."

[27]ACE: "We know the individual ended at E, which means he goth there either from C or D, C being more frequent. He can get to C from A or B, with A being more frequent. So rolling back we can say route ACE is the likely path.".

5.4 *Results*

The results of the three Z-versions are summarized in Table 11:

DG3	Z1	Z1 -distribution	Z2 - transitions	Z3 - visits
ACE	4	10% [8: 2H^{min},4T^{max}]⟨6⟩	25% [22: 4H^{sum}, 13 T^{max}]⟨8⟩	54% [57: 15 H^{sum}, 32T^{max}, 4T^{bi}]⟨33⟩
AFG	13	⟨12⟩	⟨12⟩	⟨11⟩
BCE	5	10% [8: 1T^{bi}]⟨13⟩	14% [12: 6 T^{bi}, 3 T^{rep}]⟨11⟩	17% [18: 12 T^{rep}, 3T^{bi}]⟨18⟩
BDE	6	80% [61: 53H^{max}, 3T^{rep}]⟨ 6⟩	61% [54: 32H^{max},18 T^{rep}]⟨23⟩	29%[30: 25H^{max},3T^{rep}]⟨10⟩
BDFG	1	⟨2⟩	⟨4⟩	⟨5⟩
n=		[77]⟨39 =34%⟩	[88]⟨58=40%⟩	[105]⟨77=42%⟩

Table 11: Z – results

As in the other experiments, we excluded subjects who did not provide an explanation of their choice, who mistakenly justified a different choice or who provided an explanation that made no sense. In this set of experiments, we also excluded the 10% of subjects who chose a path that does not end at E.

In Z1, almost all of the subjects (80%) chose BDE and most of them justified it using H^{max}. Only 2 subjects used H^{min} (ACE) and very few gave atomistic explanations.

In Z2, BDE remains the most popular choice, probably because both H^{max} and T^{rep} lead to it. More importantly, a slight majority of subjects used atomistic procedures. Almost all of the subjects who chose holistic procedures applied H^{max} while only 4 used H^{sum}.

In Z3, we observed an even larger decline in the proportion of subjects who chose the most frequent path, namely BDE. Both the rule of thumb H^{sum} and the atomistic procedure T^{max} became more popular and both lead to ACE.

Response Time: The response time data supports the explanation that the choice of BDE in Z2 and Z3 is an outcome of the complexity of calculating H^{max}. In Z2, the MRT of the subjects who chose BDE was 155s and the MRT of the subjects who chose ACE or BCE was 124s. In Z3, the difference is even more dramatic: BDE (251s) vs. ACE and BCE (115s).

6. Evaluating the procedures

6.1 Rules of thumb

In the experiments where the distribution of choices was not explicitly provided (V2, V3, W2 and W3) – even though they could be derived from the data – some of the procedures used by the subjects can be thought of as rules of thumb for choosing the most frequent alternative.

Evaluating a rule of thumb depends on the underlying probabilistic assumptions regarding the circumstances in which it is used. We will evaluate a rule of thumb using the probability that it leads to the most frequent alternative under the following three probabilistic assumptions:

S1: The frequencies of the 6 alternatives are drawn from a uniform distribution over the simplex.

S2 -19: The frequencies of the 6 alternatives are derived from 19 independent draws from the uniform distribution over the alternatives.

S2 -1001: The frequencies of the 6 alternatives are derived from 1001 independent draws from the uniform distribution over the alternatives.

It is well known that the frequency vector $(X_1, X_2, X_3, X_4, X_5, X_6)/\sum_{i=1,..,6} x_i$, where (X_i) is a vector of independent and exponentially distributed random variables with mean 1, is uniformly distributed over the 5-dimensional simplex. Applying this and using Wolfram Mathematica[28] enables the calculation of success for each of the rules of thumb under S1. We estimated the probabilities of success under the other two probabilistic models by running a simulation with one million iterations. For each draw, we examined the probability that the sequence that was randomly selected using a rule of thumb is a maximum of the underlying distribution. The success rates are summarized in table 12:

	Procedure	S1	S2-19	S2-1001
V2	H^{sum}	883/952≈ 0.927	0.943	0.867
W2	$T^{\#wins}$	4519/7560≈ 0.598	0.543	0.446
V3/W3	H^{sum}	7/8=0.875	0.848	0.736
V3/W3	T^{max}	3473/4320≈ 0.804	0.761	0.636

Table 12: Simulations of the rules of thumb

H^{sum} appears to be quite efficient by these measures, whereas the atomistic procedures $T^{\#wins}$ and T^{max} are less so. Another procedure mentioned by subjects in W2 is

[28]We thank Aron Tobias for suggesting and carrying out the calculations using Wolfram Mathematica.

$T^{majority}$, which leads to a cycle with probability 1/16=0.0625. Conditional on not generating a cycle, its success rate is 56/75≈0.747.

6.2 On the probabilistic rationale of atomistic procedures

One way to think about the outcomes of the V and W sets is as a realization of some probabilistic process whose parameters are unknown. Evaluating the parameters is often carried out using a maximum likelihood calculation. Arad and Gayer (2012) present an experiment in which subjects appear to evaluate the proportion of a color in an urn of colored balls according to the proportion of the color in a sample of past draws with replacement.

According to one particular model, there exists a probability measure over the alternatives (sets of size 2 in B, sequences of length 3 in V, etc.) and the data is a realization of a number of independent draws from that process. The maximum likelihood calculation estimates the probability of each alternative according to its frequency in the data. Given this model, the rational guess would be the most frequent alternative in the data.

In a different probabilistic model that can be applied to V and W, the data is created by a process built on p, a probability measure over a set $\{a,b,c\}$ (the sites A,B and C in V and the interests A,P and S in W). The first element in a chosen sequence is assumed to be the realization of a draw of p. The second is the realization of p conditional on the two remaining elements. Under this probabilistic model, the maximum likelihood estimator may assign the highest probability to an element that is never observed first. To see this, let $p(a)=\alpha$, $p(b)=\beta$ and $p(c)=\gamma$. For example, given the data $D=\{acb,bca\}$, the maximum likelihood estimator is the maximizer of $[\alpha\gamma/(1-\alpha)][\beta\gamma/(1-\beta)]$ which is $(\alpha,\beta,\gamma)=(1-\frac{1}{\sqrt{2}},1-\frac{1}{\sqrt{2}},\sqrt{2}-1)\sim(0.29,0.29,0.41)$.[29] Thus, the most likely sequence given the data D is cab or cba although c is not observed in first place within D. When applied to the data in W1, the maximum likelihood estimator is the solution to the maximization of $(a\frac{p}{1-a})^8(p\frac{a}{1-p})^5(p\frac{s}{1-p})^7$ which is $(a,p,s)=(0.24,0.66,0.11)$. Given this estimator, the best guess is PAS which is not the most frequently observed sequence. In the case of V1*, the maximum likelihood estimator assigns the probabilities $(0.26,0.39,0.35)$ to A,B,C which implies that BCA is the most likely sequence.

Note that under another probabilistic model, according to which the first element is drawn randomly and then the second is drawn randomly according to a probability measure that depends on the first realization, the maximum likelihood estimator always leads to the choice of the most frequently observed sequence.

---

[29] Setting the derivative of $log[\alpha\gamma/(1-\alpha)][\beta\gamma/(1-\beta)]$ equal to 0 gives $\frac{1}{\alpha}+\frac{1}{1-\alpha}=\frac{1}{\beta}+\frac{1}{1-\beta}=\frac{2}{\gamma}$ which implies that $\frac{1}{\alpha}+\frac{1}{1-\alpha}=\frac{1}{\alpha(1-\alpha)}=\frac{2}{1-2\alpha}$, i.e. $1-4\alpha+2\alpha^2=0$.

7. Related literature

The distinction we make between holistic and atomistic prediction procedures is somewhat related to the distinction between wide and narrow choice bracketing. When a subject needs to choose one alternative from $\{A, B\}$ and another from $\{C, D\}$, narrow bracketing means that he makes two separate choices, whereas wide bracketing means that he makes a single choice from $\{AC,\ AD,\ BC,\ BD\}$. Read, Loewenstein and Rabin (1999) demonstrated that people often fail to bracket widely even when it is beneficial for them to do so (for example, the case where A is a sure gain of $240, B is a lottery that yields a gain of $1000 with probability 0.25, C is a sure loss of $750 and D is a lottery that yields a loss of $1000 with probability 0.75).

Note that in the narrow-wide bracketing experiments, a subject has to make two separate choices and the question is whether he takes into account the link between them. In contrast, in our experiments a subject has to make one choice (a prediction) and the question is whether during the deliberation stage he "atomizes" his choice or views the choice holistically.

Another somewhat related literature investigates the procedures used to choose between two alternatives, each presented as a vector. An example would be the choice between a lottery that yields $\$x$ with probability p and a lottery that yields $\$y$ with probability q. In this context we observe holistic procedures, such as comparing the expectations of the two lotteries. But we also observe atomistic procedures that ignore a dimension in which the values are "similar" and make a decision by comparing the values in the other dimension (see, for example, Rubinstein (1988, 2013)).

The experiments in set Z which involve a directed graph were inspired by Glazer and Rubinstein (2021). In that purely theoretical paper we modeled a *story builder* as a procedure for extending partial information about an event to form a complete and coherent *story* (a path from the origin of a graph to a terminal node). That paper provided several examples of story builders – some of them holistic and others atomistic. The completed story can be thought of as an answer to the question subjects were presented with here. Note however that the two frameworks are very different. In Glazer and Rubinstein (2021), the story builder did not have the information he is provided with here. There, we focused on a consistency property (referred to as *stickiness*) of a story builder across instances in which he has different information about a particular event while here we study the construction of a story based on past events.

The differences in the results between V1 and V2; V1* and V3; W1, W2 and W3; and Z1, Z2 and Z3, can be viewed as a framing effect since the underlying distributions of choices in each group are identical. However, this is not a standard framing effect in which two presentations of the same choice problem differ in a way that affects behavior (such as presenting the data in positive or negative terms as in Tversky and Kahneman (1986)). In V2, V3, W2, W3, Z2 and Z3 subjects are required to derive the distribution of choices from the data provided, a task that is not straightforward and therefore the behavior observed in these experiments is very different than their behavior when they observe explicitly the corresponding distributions of choices.

A related framing effect is suggested by Enke and Zimmerman (2019). Some of their subjects were asked to estimate a state of nature based on four uncorrelated signals: A, B, C and D, where the optimal estimate is their average. Most of the subjects indeed seem to have chosen the average. Other subjects were asked to estimate the state of nature based on $\alpha = A$, $\beta = A + B$, $\gamma = A + C$, $\delta = A + D$. The optimal estimate in this case is the average of α, $\beta - \alpha$, $\gamma - \alpha$, and $\delta - \alpha$; nevertheless, some of the subjects still averaged over α, β, γ and δ, thus ignoring the correlation between them.

8. References

Arad, Ayala and Gabi Gayer (2012). "Imprecise Datasets as a Source of Ambiguity: A Model and Experimental Evidence". *Management Science 58*, 188-202.

Enke, Benjamin, and Florian Zimmerman (2019). "Correlation Neglect in Belief Formation". *Review of Economic Studies 86*, 313-332.

Gigerenzer, Gerd and Wolfgang Gaissmaier (2011). "Heuristic Decision Making". *The Annual Review of Psychology, 62*, 451-482.

Glazer, Jacob and Ariel Rubinstein (2021). "Story Builders". *Journal of Economic Theory 193*, article 105211.

Read, Daniel, George Loewenstein and Matthew Rabin (1999). "Choice Bracketing". *Journal of Risk and Uncertainty 19*, 171-197.

Rubinstein, Ariel (1988). "Similarity and Decision-Making Under Risk". *Journal of Economic Theory 46*, 145-153.

Rubinstein, Ariel (2007). "Instinctive and Cognitive Reasoning: A Study of Response Times". *Economic Journal, 117*, 1243-1259.

Rubinstein, Ariel (2013). "Response Time and Decision Making: A "Free" Experimental Study". *Judgement and Decision Making, 8*, 540-551.

Simonson, Itamar (1990). "The Effect of Purchase Quantity and Timing on Variety Seeking Behaviour". *Journal of Marketing Research 27*, 150-162.

Tversky, Amos and Daniel Kahneman (1971). "Belief in the law of small numbers". *Psychological Bulletin* 76, 105-110.

Tversky, Amos, and Daniel Kahneman (1986). "Rational Choice and the Framing of Decisions." *The Journal of Business 59*, S251-278.