Individual differences in base rate neglect: A fuzzy processing preference index

Christopher R. Wolfe *, Christopher R. Fisher

Miami University, Oxford, OH, USA

A R T I C L E   I N F O

Article history:
Received 30 June 2012
Received in revised form 16 January 2013
Accepted 4 March 2013

Keywords:
Individual differences
Base rate neglect
Base-rate respect
Process dissociation procedure
Fuzzy-Trace Theory

A B S T R A C T

Little is known about individual differences in integrating numeric base-rates and qualitative text in making probability estimates. Fuzzy-Trace Theory predicts a preference for fuzzy processing. We conducted six studies to develop the FPPI, a reliable and valid instrument assessing individual differences in this fuzzy processing preference. It consists of 19 probability estimation items plus 4 “M-scale” items that distinguish simple pattern matching from “base rate respect.” Cronbach’s Alpha was consistently above 0.90. Validity is suggested by significant correlations between FPPI scores and three other measurers: “Rule Based” Process Dissociation Procedure scores; the number of conjunction fallacies in joint probability estimation; and logic index scores on syllogistic reasoning. Replicating norms collected in a university study with a web-based study produced negligible differences in FPPI scores, indicating robustness. The predicted relationships between individual differences in base rate respect and both conjunction fallacies and syllogistic reasoning were partially replicated in two web-based studies.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

The extent to which people employ or ignore base rate information has been an active topic of research for about 40 years (Ajzen, 1977; Kahneman & Tversky, 1973, 1996; Reyna & Brainerd, 2008; Wolfe, 1995). Most of the research on base rate use has focused on general statements about human cognition rather than individual differences. The work presented here examines individual differences in the use of base rate information and is guided by Fuzzy-Trace Theory (FTT; Reyna & Brainerd, 1995, 2007; Reyna, 2012). Our goal is to create an index that reliably assesses the extent to which individuals respect or neglect base rate information in making probability estimates.

Theoretically, our work is guided by FTT (Reyna, 2012) a dual process theory. Like other dual process theories, FTT is compatible with the notion that there are radical redundancies in our cognitive architecture. Higher-order thinking can be accomplished through more than one mechanism or process operating on more than one kind of representation. FTT holds that when people encode information they create multiple representations from precise verbatim representations of surface characteristics at one end of the continuum, to vague gist representations encoding the essential bottom line meaning of events at the other end of the continuum (Reyna & Brainerd, 1995). The terms gist and verbatim are used much as they are in everyday language to capture the distinction between the exact detail or wording and the underlying meaning. A key provision of FTT is that people have a preference to reason with the vaguest most gist-like representation permissible for any problem. This is known as the fuzzy-processing preference (Reyna & Brainerd, 2011).

FTT has a kinship with other dual process theories, and there are similarities between the gist-processing and verbatim–processing distinction and System 1 and System 2, (Kahneman, 2011; Stanovich & West, 2000) associative and rule-based processing, (Sloman, 1996) heuristic and rule-based, (Ferreria, Garcia-Marques, Sherman, & Sherman, 2006) and heuristic and analytic processing (Evans, 2008). However, there are also important differences between FTT and other dual process theories. Novel predictions of FTT that have been confirmed empirically include a developmental trend toward increasing gist-processing with age (Reyna & Casillas, 2009), people with autism rely more on verbatim processing and less on gist based reasoning and more on verbatim-based, (Reyna & Brainerd, 2011) and experts with a good deal of domain knowledge exhibit more gist processing than novices (Reyna & Lloyd, 2006). Gist processing allows experts to make sharper, more meaningful distinctions (Reyna & Lloyd, 2006), and reduce overprecision errors in judgment (Haran, Moore, & Morewedge, 2010). Thus, according to FTT, gist processing is not restricted to casual decision making when the stakes are low. Rather, gist processing is the source of both mature expert performance and systematic errors.

FTT has illuminated the cognitive processes underlying several cognitive illusions in the judgment and decision-making literature (Reyna & Adam, 2003). For example, the conjunction fallacy, displayed when people erroneously estimate P(A and B) > P(A), is best understood as a problem of reasoning with nested sets (Reyna & Brainerd, 2011). Joint probability problems are confusing because people must compare relevant denominators as well as numerators. To illustrate, consider a version of the famous Linda problem (Wolfe & Reyna, 2010); Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student she was deeply concerned with issues of discrimination and
social justice, and also participated in anti-nuclear demonstrations. She lives in a cooperative house called Women’s Space. What is the probability that Linda is a bank teller? What is the probability that Linda is a feminist? What is the probability that Linda is a bank teller and a feminist? Many people will fallaciously provide a greater probability estimate for, “What is the probability that Linda is a bank teller and a feminist?” than for “What is the probability that Linda is a bank teller?” even if both questions are presented on the same computer screen at the same time (Wolfe & Reyna, 2010). Unfortunately, people frequently base probability judgments solely on comparisons between numerators, identified in FTT as denominator neglect (i.e. the set of feminist bank tellers is a numerator that can be compared to the denominator feminists and also the denominator bank tellers). This oversimplifies the problem, allowing people to lose track of the relationships between the numerators and focus only on the denominators in which they are included (Wolfe & Reyna, 2010).

Dual process theories offer a promising avenue to investigate individual differences in cognition. If people are capable of employing more than one type of process or representation for a particular task, then it is reasonable to ask whether individuals might reliably differ in their tendency to employ one sort of process over another in a given context. In the current work, we hypothesize that, within the framework of a general fuzzy processing preference, people have reliable preferences for more gist-based or verbatim-based judgments in integrating numeric base rates with qualitative verbal information. These hypotheses will be spelled out more fully with an example below.

One promising candidate for examining individual differences in the use of base rates is the Process Dissociation Procedure (PDP; Jacoby, Lindsay, & Toth, 1992; Ferriena et al., 2006). The PDP was originally developed to assess unconscious influences on attention, awareness, and control (Jacoby et al., 1992). Since its inception, PDP has been successfully applied in the domain of judgment and decision making to test central tenets of dual process theories (Ferriena et al., 2006). One useful property of PDP is that parameters can be estimated at the level of the individual and thereby serve as an index of individual differences.

To illustrate the use of PDP for base rate problems, consider the following problem used in the current work, designed to portray Bob as college-bound: “At Central High School 10% of the seniors go on to college. Bob is a senior at Central High. He gets mostly As and Bs in school and is well liked by his teachers. Which is more likely, (a) Bob will go to college or (b) Bob will not go to college?” There are two components of this base rate problem: the stated base rate and the text description. The general PDP procedure requires the presentation of two problem versions: an “inclusive” version in which the stated base rate is congruent with the text description, (e.g. At Central High School 90% of the seniors go on to college) and an “exclusive” version in which the stated base rate is incongruent with the text description (e.g. the version above where 10% of the seniors go on to college.) For each version, participants make a forced choice. Following Ferriena et al. (2006) participants who favor a heuristic approach should lean on the text description whereas those who favor a rule-based approach should more heavily weight the specifically stated numeric base rates.

The relative contribution of rule-based and heuristic processes for a given individual are computed according to the following formula (Ferriena et al., 2006): 1

\[
\text{Rule Based} = \frac{P(\text{inclusive version is choice a}) - P(\text{exclusive version is choice a})}{2P(\text{exclusive version is choice a})}
\]

\[
\text{Heuristic} = \frac{P(\text{exclusive version is choice a})}{2} - \text{Rule Based}
\]

An important assumption of the PDP – one to which we shall return – is that the processes are independent.

The PDP is useful for many purposes, but we argue that with respect to base rate, it can be improved upon for theoretical and methodological reasons. Thus, we developed and tested an alternative instrument, the Fuzzy Processing Preference Index (FPPI). Our overarching goal is to assess individual differences in a preference for gist based estimates (text) and verbatim based estimates (base rate) in problems where quantitative base rates and qualitative descriptions are integrated to form probability estimates. For participants unschooled in statistical reasoning, the gist-based tendency is to heavily weigh the text description whereas the verbatim-based tendency is to more heavily weight the specific numeric base rate.

A gist representation is one that captures the essential characteristics of a situation. In the case of the Bob problem described above, the bottom-line meaning that most people (unfamiliar with the base rate concept) derive is that Bob is a good student. For many people, this is consistent with their representation of a college bound high school student. However, attention to a specific detail, that 10% of the students at Bob’s school go on to college, reveals a discrepancy that must be included in the assessment of the chances that Bob will go to college. Theoretically stated, base rate problems present us with a number of nested hierarchical relationships, only some of which are known. Thus in the Bob problem the base rate for college matriculation from Bob’s school is 10% and the rate at which other students getting As and Bs is unknown. It can thus be considered a form of conditional probability problem, i.e. what is the probability that Bob will go to college given that 10% of the students from his school go to college. As with other conditional probability problems, FTT suggests that even people who understanding the meaning of the base rate are likely to consider them peripheral because they erroneously reason that the base rate is less important than specific information (Wolfe, 1995). Thinking this way, getting As and Bs is specifically about the topic of interest (Bob), while the base rate is “merely” about other students. However, a person who is more inclined to attend to verbatim details is more likely to weigh the base rate in probability judgments, perhaps through a process of anchoring and adjustment. These probability problems, with their layers of nested and overlapping inclusion relationships, make for complicated mental bookkeeping, and “denominator neglect” is a common simplification strategy. Antidottally, this is captured in a remark by one student, “why should we care about other people when the story is about him.” From an FTT perspective, base rate neglect is more than a difficulty in reasoning with percentages, and materials using frequencies (Gigerenzer & Hoffrage, 1995) or verbal labels would likely produce comparable results.

In the current work, we went beyond developing a procedure and created a valid and reliable instrument that is amendable to item analysis, and not reliant on forced choice. Moreover, our instrument requires less time for administration. Following FTT we do not make an assumption that rule based and heuristic processes are independent. Rather we view processing preferences along a continuum from vague, fuzzy gist representations to more precise verbatim representations (Reyna & Brainerd, 1995). People having a general preference for gist processing (Reyna, 2012; Reyna & Brainerd, 1995) and in this research we investigate whether individuals reliably differ in their preferences for verbatim base rate information along this continuum. A central goal of this research is to create a useful and valid approach to reliably assessing individual and experimental group differences in how people understand and make probability estimates. Thus, we developed a more sensitive individual differences measure than the PDP, with better psychometric properties.

The formula for the FPPI is:

\[
V = \frac{|1 - |E - B||x|E - t|}{|B - t|}
\]

where E is the estimate provided by participants; B is the base rate (given value in the problem); t is the item mean for text alone.

---

3. Study 1: an assessment of the t and Bi parameters

Having developed the 19-item FPPI instrument, the purpose of the first study was to determine two methodological questions about the materials and parameters in the formula. The PDP makes use of both inclusive and exclusive base rates. Thus, the first question is whether inclusive base rates would add to the reliability of the instrument. The second question is whether the t parameter for the text without a stated base rate needs to be calculated on a per participant, per item basis, ($t_p$) or whether it best considered an item characteristic that can be calculated as a sample mean for each item ($t_m$). To illustrate, in the case of the problem about the student Bob, the first question is whether the problem needs to be presented twice, once for a school where only 10% of the students go on to college, and a second time for a school where 90% go on to college. The second is whether it is necessary to obtain estimates $t_p$ for each participant without a stated base rate as, “Bob is a senior at Cloverdale High. He gets mostly As and Bs in school and is well liked by his teachers. What is the probability that Bob will go to college?”

The variations on the formula can be expressed as:

$$V = \frac{1 - |O_i - B_i| - |O_e - B_e|\times |O_e - t|}{|B_e - t|}$$

where $O_i$ is the observed probability estimate made by participants in the inclusive condition (i.e. stated base rate and text lead to similar probability estimates); $B_i$ is the stated inclusive base rate (i.e. 90% for the Bob problem); $O_e$ is the observed probability estimate made by participants in the exclusive condition (i.e. stated base rate and text lead to opposite probability estimates); $B_e$ is the stated exclusive base rate (i.e. 10% for the Bob problem); and $t$ is the estimate for the text without a stated base rate calculated either $t_p$ on a per participant per item basis or $t_m$ as the mean of all estimates for an item from entire sample of participants.
of the FPPI. Cronbach’s Alpha was calculated for the 19 items two ways, for $t_p$ and $t_m$. For $t_m$, Cronbach’s Alpha = .91, and for $t_p$, Cronbach’s Alpha = .87. This indicates that with respect to reliability, we found no advantage to collecting estimates without stated base rates for each participant. The $t_m$ parameter estimates based on sample means for this study ($n = 96$) are provided in parentheses for each item in Appendix A.

Cronbach’s Alpha was also calculated for the 19 items with and without the inclusive estimates. For the exclusive estimates alone Cronbach’s Alpha = .91. For the inclusive and exclusive estimates Cronbach’s Alpha = .87. We found no advantage to using the inclusive condition with respect to reliability.

This study employed a fairly laborious procedure with similar probability estimation tasks repeated three times and separated by filler tasks. Taken collectively, these results suggest that such a procedure is unnecessary. Simply administering the FPPI items once with exclusive base rates yielded a solid Cronbach’s Alpha = .91 that was no less reliable than with the more time consuming and elaborate procedure. However, the comparison of inclusive and exclusive condition to the inclusive condition alone was replicated in the following study.

4. Study 2 PDP vs. FPPI

This study is a direct comparison of the FPPI with the PDP procedure using the same set of 19 items. Because the constructs of gist and verbatim processing are related to the constructs heuristic and rule-based processing, the general predictions are that the FPPI will positively correlate with rule-based and negatively correlate with heuristic scores, while providing superior psychometric properties.

4.1. Method

4.1.1. Participants

There were 148 undergraduates who participated individually or in groups of 2–4 in the laboratory at individual computers and received credit toward a requirement in an introductory psychology course at Miami University.

4.1.2. Materials

The problems were the 19 FPPI items and 6 filler items (see Appendix A). Problems were presented in two forms, FPPI and PDP. In the FPPI conditions items required a probability estimate as in the previous studies. We included both inclusive and exclusive conditions for both the PDP and FPPI problems. In the PDP conditions the problems required a forced choice decision. For example, “At Central High School (Inclusive = 90% or Exclusive = 10%) of the seniors go on to college. Bob is a senior at Central High. He gets mostly As and Bs in school and is well liked by his teachers. Which is more likely, (a) Bob will go to college or (b) Bob will not go to college?” The $t$ parameter for text alone was estimated using the sample mean from Study 1.

4.1.3. Design and procedure

The study was substantially a within subjects design with task order and list counterbalanced in a $2 \times 2$ Latin Square and treated as a nuisance variable. Table 1 shows the four counterbalanced task orders. There were no significant differences for the between subjects manipulations, $F(3, 144) = 1.08, p = .36$, and all subsequent analyses were conducted on a within participants basis.

Participants completed tasks in one of the four task orders depicted in Table 1 after providing their informed consent. After each substantive task, participants completed a 20 minute filler task on argumentation. Thus, there were four substantive tasks and three filler tasks. Upon completing the study participants were thanked and debriefed.

4.2. Results and discussion

For the PDP inclusive conditions (e.g. Bob College 90%), participants had a mean of 18.24 of 19 problems (96%) selected in the affirmative and in the exclusive condition they had a mean of 11.45 of 19 (60%) affirmative. This produced a Rule Based Mean of 0.36 (SD = .32) and Heuristic mean of 0.90 (SD = .22). The correlation between the Rule Bases and Heuristic scores was $- .37$.

Turning to the FPPI, using only the exclusive condition the FPPI had a Cronbach’s Alpha = .94 and using both inclusive and exclusive estimates Cronbach’s Alpha = .96. This suggests that the procedure is quite reliable even without the addition of the more cumbersome inclusive condition. Thus in all subsequent analyses we used only the data from exclusive version.

The FPPI had a mean of .47 (SD = .28). Comparing the two approaches the correlation between the FPPI and PDP Rule Based scores was $r = .69$ and the correlation between the FPPI and Heuristic scores was $r = -.46$. The high Cronbach’s Alpha of .94 provides more evidence for very good item reliability even in the absence of the more time consuming inclusive condition. The high correlation between FPPI and PDP rule based scores suggests that they are measuring similar constructs. This can be conceived as evidence of task validity, the focus of the next study.

5. Study 3: using the FPPI to predict conjunction fallacies and logic in syllogistic reasoning

To the extent that the FPPI measures individual differences in base rate respect, the instrument should shed light on individual differences in other cognitive processes making use of gist and verbatim representations. This study examined the relationship between the FPPI, conjunction fallacies in joint probability estimation, and logic in syllogistic reasoning.

The essence of fallacies in joint probability estimation is erroneously ascribing a higher probability to a joint probability than to one of the constituent probabilities (Wolfe, Fisher & Reyna, 2012). FFT explains conjunction fallacies as both errors of processing and errors of representation. Specifically, such errors are attributed to denominator neglect—behaving as if one is ignoring the relevant denominators in a $2 \times 2$ table (Wolfe & Reyna, 2010). This leads to the clear prediction that the more an individual scores as high in a preference for gist processing (i.e. the lower the FPPI score) the more likely that person will be to commit the conjunction fallacy in joint probability estimation.

A more distant cognitive domain is syllogistic reasoning in which participants must assess the structure of the argument for validity while ignoring the believability of its semantic content. Making valid inferences with syllogisms does not require the use of base rates, but it does require careful attention to the verbatim details of the problem, especially the key connective terms such as “some,” “all,” “no,” and “not.” This is particularly important when beliefs and logic come into conflict. Consider the following syllogism, (Evans, Handley, & Harper, 2001) “No addictive things are inexpensive. Some cigarettes are inexpensive. Therefore, some addictive things are not cigarettes.” The conclusion is believable based on world knowledge. However, it is not logically valid. By way of contrast, the following is a valid syllogism that does not have a believable conclusion, (Evans, Barston, & Pollard, 1974).

<table>
<thead>
<tr>
<th>Task Order</th>
<th>FPPI list A</th>
<th>PDP list A</th>
<th>FPPI list B</th>
<th>PDP list B</th>
</tr>
</thead>
<tbody>
<tr>
<td>FPPI list A</td>
<td>FPPI list A</td>
<td>PDP list A</td>
<td>FPPI list B</td>
<td>PDP list B</td>
</tr>
<tr>
<td>FPPI list B</td>
<td>FPPI list B</td>
<td>PDP list B</td>
<td>FPPI list A</td>
<td>PDP list A</td>
</tr>
</tbody>
</table>

Table 1

Counterbalancing of task order.
ences judgments of logical validity (Reyna & Brainerd, 2011). Thus, as valid and believable, invalid and believable, valid and not believable, and invalid and not believable Evans and Curtis-Holmes (2005) developed an index to the extent to which people solve syllogisms logically (the logic index) and the extent to which they are subject to belief bias (the belief bias index). FTT posits that gist representations of syllogisms include content or meaning, rather than being purely abstract, accounting for belief bias in which factual knowledge influences judgments of logical validity (Reyna & Brainerd, 2011). Thus, as with probability judgments, syllogisms are subject to interference from overlapping classes or class-inclusion confusion. Our prediction for this more remote cognitive domain of syllogistic reasoning is that the preferential use of verbatim representations would be associated with logical responding.

5.1. Method

5.1.1. Participants

There were 124 undergraduate introductory psychology students at Miami University who participated individually or in groups of 2–4 at individual computers in the laboratory to fulfill a course requirement.

5.1.2. Materials

Materials were the FPPI (external items only) used in previous studies, joint probability estimation task, and a syllogistic reasoning task. For the FPPI, the t parameter for text without stated base rates was calculated as the mean estimate from Study 1.

The joint probability task consisted of 12 problems such as the Linda problem described in Section 1. For each problem participants read a short description and then made 4 probability estimates P(A), P(B), P(A and B), and P(A or B) on the same screen. Each problem was scored as a conjunction fallacy if the estimate of P(A and B) > P(B). Thus conjunction fallacy scores could thus range from 0 to 12 (if someone committed a conjunction fallacy on each of the 12 problems).

We used a total of 16 syllogisms that were created by factorially varying the believability of the conclusion (believable vs. unbelievable) and the validity of the syllogism (valid vs. invalid), with four of each type. The instructions, and several syllogisms, were taken from Morley, Evans, and Handley (2004) with additional syllogisms taken from Cherubini, Garnham, Oakhill, and Morley (1998) and Dube, Rotello, and Heit (2010) plus three original syllogisms. Thus we have solid support from the literature that the syllogisms match our perceptions of believability and validity. Each syllogism that was rated as valid was scored as a 1 and each that was rated invalid was scored as a 0. The logic index was calculated as the sum of valid-believable items plus valid-unbelievable items minus the sum of invalid-believable items plus invalid-unbelievable items. Thus a participant who rated each valid syllogism as valid and each invalid syllogism as invalid would receive a logic index score of 8. The belief index was calculated as the sum of valid-believable items plus invalid-unbelievable items minus the sum of valid-unbelievable items plus invalid-believable items. Thus a participant who rated each believable syllogism as valid and each unbelievable syllogism as invalid would receive a belief index score of 8.

5.1.3. Design and procedure

Once again the study substantively employed a within subjects design with task order manipulated between participants as a nuisance variable. After providing their informed consent, all participants received the FPPI first. Half of the participants were randomly assigned to receive the syllogistic reasoning task second and the joint probability estimation task third while the remaining participants received the joint probability task second and the syllogistic reasoning task third. There was no significant effect for task order. When participants completed all tasks they were thanked and debriefed.

5.2. Results and discussion

The mean score on the FPPI was .57 (SD = .27) and Cronbach’s Alpha = .93. This mean was considerably higher than in previous studies, which is partially attributable to 17 participants who used the stated base rate as their estimate on 18 or 19 of the 19 items. This study was conducted during the last three weeks of the academic semester so perhaps some participants were rushing through the study. This finding underscores the need to include a mechanism that distinguishes between participants who are high in base rate respect, and those who employ a simple matching strategy of using the stated estimate in the problem. Toward this end we developed a set of M-scale items to distinguish between base rate respect and a simple matching strategy. The M-scale was tested in the next three studies. In the current study we analyzed the data two ways, once with all participants and one excluding those participants who used the stated base rate on 18 or 19 of the 19 items. With the reduced set of 107 participants the mean score on the FPPI was .51 (SD = .24) and Cronbach’s Alpha = .90.

Turning to conjunction fallacies on the joint probability estimation task, there was a significant negative correlation between the FPPI and the number of conjunction fallacies. The more participants scored toward the gist end of the scale the more they committed the conjunction fallacy of estimating P(A and B) > P(B). For all 124 participants r = −.28, p < .002. For the reduced set of 107 participants r = −.32, p < .001. FTT suggests that these conjunction fallacies are produced by denominator neglect, which is a result of gist processing. Thus, we found the predicted association between base rate neglect and conjunction fallacies.

For some purposes it may be useful to have a cutoff score to distinguish between people at the gist end and those at the verbatim end of the FPPI continuum. Although FTT conceives the verbatim-gist distinction as a continuum, as with other dual process theories, researchers sometimes simplify to a dichotomy of gist and verbatim representations. Moreover, for some research purposes such as assigning participants to research conditions and bringing conceptual clarity to statistical interactions it is useful to have a trustworthy dichotomized variable. We found that a score of FPPI = .45 reliably distinguishes between the number of conjunction fallacies. Using data for all participants, 73 scored toward the verbatim end (FPPI > .45) and had a mean of 4.66 conjunction fallacies, and 51 scored toward the gist end (FPPI < .45) and had a mean of 6.22 conjunction fallacies, F(1,122) = 9.55, p < .003. Using the reduced set of participants, 56 scored toward the verbatim end (FPPI > .45) and had a mean of 4.63 conjunction fallacies, and 51 scored toward the verbatim end (FPPI < .45) and had a mean of 6.22 conjunction fallacies, F(1,105) = 8.87, p < .004. We will test this cutoff score in subsequent studies.

Considering the more distant task of syllogistic reasoning, with all 124 participants the correlation between the FPPI and the logic index approached statistical significance with r = .16, p = .06. With the reduced set of 107 participants the correlation between the FPPI and the logic index score was r = .18, p < .05. The correlation between the belief index and FPPI was not significant. For all 124 participants r = −.09, p = .31 and with the reduced set of 107 participants r = −.16, p = .11.

The large number of participants who used the stated base rate on all or nearly all of their responses highlights the need for a technique to distinguish between a high degree of base rate respect and a simple mapping strategy. Thus, the next three studies were used to develop and test a set of M-scale items. We also replicated the substantive findings reported above concerning conjunction fallacies and the logic index on the syllogistic reasoning task.
6. Study 4: norm verification and M-scale development

The previous studies suggest that we could reliably estimate the t parameter representing the probability estimates yielded by the text without a stated base rate by using sample means from Study 1. Because this t parameter plays an important role in the FPPI, we wanted to assess the reliability and robustness of norms based on those sample means. For this reason, we sought to replicate the first portion of Study 1 with a diverse sample by asking for probability estimates for each problem without and stated base rates. To collect data from a diverse sample, we conducted this study on the web using the European-based Wextor platform (Reips, 2000).

To distinguish between a healthy degree of base rate respect and a simple matching strategy of parroting the stated base rate as the probability estimate we created a set of M-scale items. The ideal for these M-scale items was to create items that could plausibly have a very high or very low stated base rate, but for which the particular details stated verbally in the text would make the probability of the event almost impossible or almost certain and in the opposite direction of the stated base rate. Here is a sample item, and all of the items finally included in the M-scale are included in Appendix B. “Richard is an avid skier and spends 90% of his vacations skiing. Today he has plane tickets to Aspen, Colorado and has been looking forward to this weekend trip for months. Unfortunately, Richard had a bad accident and both of his legs are broken. What is the probability that Richard will go skiing this weekend?” It is highly unlikely that Richard could go skiing with two broken legs. We reasoned that anyone who estimates the probability that Richard will go skiing this weekend as 90% is employing a matching strategy. Participants who gave the stated probability as their estimates for any of our M-scale items are considered to be using a matching strategy rather than displaying high levels of base rate respect. Thus, their data can reasonably be excluded from standard analyses and treated separately.

6.1. Method


6.1.1. Participants

People participated in this web-based study on the basis of interest without any compensation. Participants were free to leave the study at any time, and skip any items they choose to skip. There were 295 participants who answered at least one question including demographic questions, and 151 participants who made estimates for all of the items. Demographic data was collected as participant self-reports. Of the 292 participants who provided data on age and gender, the mean age was 24.3 years old (SD = 11.5) and 129 (44%) were male, 161 (55%) were female, and 2 reported “other.” Of the 151 participants who made estimates on all of the problems, 150 answered questions about age and gender, the mean age was 26.2 years old (SD = 12.0) and 69 (46%) were male, 80 (53%) were female, and 1 reported “other.”

Table 2 shows the self-reported location of participants by content, indicating that about 20% of the participants who completed all of the items were located outside North America. These data suggest that this sample is different from the Miami University sample used in Study 1.

6.1.2. Materials

The materials were the 19 FPPI items and they were presented one at a time in random order. We also included 5 candidate M-scale items with a stated base rate. Prior to this study, we vetted five M-scale items with 30 pilot participants and all agreed that the events described in the problem were opposite of the stated base rate — i.e. highly unlikely for M-scale problems with high base rates and highly likely for M-scale problems with low base rates. The M-scale items are presented in Appendix B. We also tested an additional item about a fictitious actress, Brittany Bright who won an Academy Award and an accountant who had “insider information.”

6.1.3. Procedure

Participants navigated to the study on the Wextor site with a standard web browser, read a brief task description, and clicked on a button indicating their informed consent. Participants answered the demographic questions and received the items one at a time in a random order. Participants were free to drop out at any time and skip over any item. Those who completed the study received a brief description and a statement thanking them for participation.

6.2. Results and discussion

Table 3 presents the number and percentage of participants who responded with the stated base rate (e.g. probability of Richard skiing is 90%) on each of the M-scale items. The majority of participants gave 0% as their estimates (or 100% for the soccer item). Only estimates that exactly matched the stated base rate were counted as M-scale “hits.” It can be seen that the item about Brittany Bright (who won an Academy Award and an accountant who had insider information about the results) produced many more base rate responses than the other M-scale items. Perhaps this is because some participants do not know that the Academy Awards hires an accounting firm to tally the results. In any case, given the discrepancy, we decided to exclude this item from future studies and keep the four M-scale items found in Appendix B. Overall, 5.1% of participants who answered at least one item and 7.3% of participants who made estimates for all of the items answered with the stated base rate on at least one M-scale item. In future work, these participants would be excluded from the analyses and their data would be treated separately.

Turning to The FPPI, Table 4 reports the Mean for each item that can be used as the t parameter estimate using both the entire sample and the those participants who provided estimates for all of the items. The first column represents the name of a character found in each problem (see Appendix A) and column 2 presents the means found in Study 2 and used in the other studies to estimate the t parameter. It can be seen that the differences among means by item range from 0.85 to 12.07.

Of course, the raw means are not of interest in and of themselves. Rather, what is of interest is the extent to which they change FPPI scores. To make this assessment we recalculated the data from Study 4 using the Wextor study means to estimate the t parameter. This gave us two FPPI scores for each participant; one calculated with the Miami University means, and one calculated using the Wextor means. For each participant, we took the absolute value of difference between the two FPPI scores and the mean |Wextor FPPI – University FPPI| = 0.0055. This indicates that the FPPI scores, which range from 0 to 1, remained essentially unchanged using a completely different sample to estimate the t parameter. Not surprisingly, there was a

<table>
<thead>
<tr>
<th>Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Any items</td>
</tr>
<tr>
<td>All items</td>
</tr>
</tbody>
</table>
very high correlation between the FPPI scores calculated these two ways, \( r = .9996 \). These data suggest that FPPI is highly robust to small differences in estimates of the \( t \) parameter from two diverse samples, and that using the Miami University norms to estimate the \( t \) parameter is justified.

7. Study 5: a web-based study of the FPPI and conjunction fallacies

The purpose of this study was to replicate the finding that the greater the degree to which a person scores toward the gist end of the FPPI the more frequently they will erroneously estimate \( P(A \text{ and } B) > P(B) \). A second purpose was to study the effectiveness of the M-scale.

7.1. Method

This study was conducted on the web using the Prezza Checkbox platform housed at Miami University. However, Miami students were not invited to participate. Instead we “advertised” the web-based study on Psychological Research on the Net, (http://psych.hanover.edu/research/exponnet.html) Facebook, and WebExperiment.net (http://webexperiment.net/).

7.1.1. Participants

People participated in this web-based study on the basis of interest without any compensation. Participants were free to leave the study at any time. They could skip demographic questions, however, they could not choose to skip probability estimation items. There were 50 participants, 45 reporting that they were located in North America, 4 in Europe, and 1 in Asia. There were 20 male participants, 29 were female, and 1 reported gender at “other.”

7.1.2. Materials and procedure

We gave participants the 19 FPPI items (see Appendix A) mixed with the 4 M-scale items (see Appendix B) first and the 12 joint probability estimation problems used in Study 3 (Section 5.1.2) second. Participants received these items in one of two randomly determined orders. Participants navigated to the web-based study with a standard web browser, read a brief task description, and clicked on a button indicating their informed consent. Participants answered the demographic questions and received the items one at a time. If participants dropped out before the study was completed their data were not recorded. Those who completed the study received a brief description and a statement thanking them for participation.

7.2. Results and discussion

The mean FPPI score was .44 (SD = .26) and Cronbach’s Alpha = .93 for the 19 FPPI items. Three (3) participants answered one or more of the M-scale items with the stated base rate. One did so with the Jessica problem, 2 on the Richard problem, and 2 on the soccer problem. These participants were significantly more likely to use the stated base rate on the 19 FPPI items than the other participants. The 3 M-scale participants used the stated base rate for their probability estimate a mean of 11.67 of 19 times compared to a mean of 4.51 for the other 47 participants, \( F(1,48) = 5.54, p < .02 \). This suggests that the M-scale participants were using a simple matching strategy.

Once again we found the predicted negative correlation between FPPI scores and the number of conjunction fallacies on the joint probability estimation task. For the sample as a whole, the correlation between the FPPI and conjunction fallacies was \( r = -.46, p < .0007 \), and excluding the 3 M-scale participants \( r = -.45, p < .002 \).

In Study 3 we set FPPI = .45 as a cutoff for distinguishing between participants at the verbatim end of the FPPI continuum and those at the gist end. Using this pre-established cutoff we found that 19 participants at the verbatim end (FPPI > .45) committed significantly fewer conjunction fallacies with a mean of 1.89 compared to the 31 participants at the gist end of the continuum with a mean of 3.77 conjunction fallacies, \( F(1,48) = 6.77, p < .01 \). Excluding the 3 M-scale participants produced similar results. The 17 participants with FPPI > .45 had a mean of 2.00 conjunction fallacies compared to a mean of 3.67 for the 30 participants with FPPI < .45, \( F(1,45) = 4.83, p < .03 \). This suggests that FPPI = .45 is a reasonably good cutoff for distinguishing between individuals with a fuzzy processing preference for gist or text-based estimates and those with a clear and consistent preference for estimates based on the verbatim statement of the base rate.

Once again, the FPPI was found to be highly reliable as measured by Cronbach’s Alpha. For a second time we found that the FPPI predicts conjunction fallacies. We also found evidence that the M-Scale appropriately excluded participants who used a simple matching strategy.

8. Study 6: a web-based study of the FPPI and syllogistic reasoning

The purpose of this study was to replicate the finding that those scoring high verbatim on the FPPI are better able to identify valid syllogisms even when logic and beliefs contradict one another. A second purpose was to further test the M-scale as a technique to identify a simple matching strategy.
8.1. Method

This study was conducted on the web using the Prezza Checkbox platform advertising the study in the same was as Study 5 (Section 7.1).

8.1.1. Participants

People participated in this web-based study on the basis of interest without any compensation. Participants were free to leave the study at any time. They could skip demographic questions, however, they could not choose to skip FPPI items or syllogisms. There were 79 participants, 68 reporting that they were located in North America, 8 in Europe, 1 in Asia, 1 in Australia, and 1 reporting “other.” There were 25 male participants, 53 were female, and 1 reported gender at “other.”

8.1.2. Materials and procedure

Participants first received the 19 FPPI items (see Appendix A) mixed with the 4 M-scale items (see Appendix B) and then they got the 16 syllogism problems used in Study 3 (Section 5.1.2). Logic index and belief index scores were calculated using the same procedure described in Study 3 (Section 5.1.2) Participants received these items in one of two randomly determined orders. Participants completed the web-based study with a standard web browser, read a brief task description, and clicked on a button indicating their informed consent. Participants answered the demographic questions and received the items one at a time. If participants dropped out before the study was completed their data were not recorded. Those who completed the study received a brief description and a statement thanking them for participation.

8.2. Results and discussion

Participants had a mean FPPI score of .40 (SD = .23) and Cronbach’s Alpha = .96. Seven (7) participants used the stated base rate as their estimate on one or more of the M-scale items (89%). Two used the base rate on the Mrs. Addison problem, 2 on the Jessica problem, 1 on the Richard problem, and 3 on the Soccer problem. As in Study 5, M-scale participants used the stated base rate as their estimates significantly more often than the other participants. Of the 19 FPPI items, the 7 M-scale participants used the base rate a mean of 10.7 times compared to a mean of 2.97 for the other 72 participants, F(1,77) = 20.59, p < .0001. This suggests that participants identified by the M-scale are using a simple matching strategy.

The correlations between the FPPI and the logic index were not significant. Excluding the M-scale participants, the correlation between the logic index and FPPI was r = .15, p = .22, and the correlation between the FPPI and the belief index was r = -.05, p = .67.

However, using the pre-established cutoff of FPPI = .45, we found that the FPPI predicted both logic index and belief index scores. Using all participants, there were 24 participants at the verbatim end of the fuzzy processing preference continuum (FPPI > .45) with a mean logic index score of 3.50 and 55 participants at the gist end (FPPI < .45) with a mean logic score of 2.33, F(1,77) = 4.29, p < .042. These same verbatim participants had a mean belief index score of 1.42 and the gist end participants had a mean belief index score of 2.55, F(1,77) = 4.61, p < .035. Excluding the 7 M-scale participants leaves an n = 72. For these participants, there were 19 participants at the verbatim end of the FPPI continuum (FPPI > .45) with a mean logic index score of 3.53 and 53 participants at the gist end (FPPI < .45) with a mean logic score of 2.28, F(1,70) = 3.85, p < .05. These verbatim end participants had a mean belief index score of 1.21 and the gist end participants had a mean belief index score of 2.55, F(1,70) = 5.23, p < .025.

Once again the FPPI showed high reliability and the M-scale appropriately excluded participants who used a simple matching strategy. We also have some evidence that the FPPI predicts differences in syllogistic reasoning with those scoring at the verbatim end (FPPI > .45) having the predicted higher scores on the logic index, and those scoring at the gist end (FPPI < .45) scoring higher on the belief index. It is worth bearing in mind that estimating probabilities by integrating base rate and text is cognitively distinct from determining whether syllogisms are logically valid.

9. Norms from combined studies

Five of the 6 studies used the FPPI in its final form (Study 4 presented the items without any stated base rates). Combining data from all of these studies (n = 457) provides some insights into the way individual differences in preferences for qualitative text-based gist and quantitative verbatim base rate information are distributed. Fig. 1 shows the distribution of FPPI scores collapsed into 7 categories with a range of .14 to reveal the basic shape of the distribution. As predicted by FTT, most people score at the gist end of the continuum. The distribution skews toward the verbatim end (high).

Table 5 provides empirical norms for the FPPI combining all 5 studies (n = 457). In increments in .1, it shows the percentage of participants below a FPPI score. In addition, the 25th, 50th, and 75th percentile are highlighted along with the cutoff score of FPPI = .45. Overall, half of the participants score below FPPI = .38, and 62% score below the empirically derived and replicated cutoff score of FPPI = .45. Our data suggest that people scoring above FPPI = .45 are likely to display fewer conjunction fallacies in their joint probability estimates. Thus, in conjunction with the M-scale, FPPI = .45 is a reliable cutoff for identifying people a clear and consistent preference for estimates based on the verbatim statement of the base rate. Of course, treating the FPPI as a dichotomous variable eliminates valuable data from the continuous FPPI, and the practice of using the cutoff score of FPPI = .45 should be used with caution.

10. Conclusions

Our instrument consists of just 23 items, 19 FPPI items and 4 M-scale items (see Appendices A and B). It is easy to administer and takes less than 10 min to complete. We have developed and tested the instrument with undergraduate students and with diverse web-based samples. The instrument is consistently very reliable, it consistently predicts behavior in estimating joint probabilities.

One shortcoming of this series of studies is that we have not established the extent to which fuzzy processing preferences for respecting or neglecting numeric base rates are best considered highly stable traits or more ephemeral states. Longitudinal studies will have to be conducted to make this determination. Theoretically, we believe that such preferences are learned, and thus amenable to the influence of pedagogic interventions. Thus we would expect some stability over
time that could, none-the-less show movement depending on educational and life experiences.

An important caveat in interpreting the current work is that most participants can be presumed to be novices on the tasks at hand — making statistical judgments and assessing the validity of syllogisms. FTT suggests that experts would have developed appropriate gist for the relevant tasks, and solve the problems using superior gist processing. Thus, those who are well trained in statistical judgment have developed a gist understanding of the meaning of “base rate.” A good logician does not need to process the verbatim details of each word in a syllogism. Rather, with a gist understanding of the bottom-line meaning of logical concepts such as the fallacy of the undistributed middle, *modus tollens*, and the like, he or she focuses on the key connectives such as “some” and “all” and ignores the rest. For un schooled participants such as ours, high scores on the logic index are associated with scores at the verbatim end of the FPPPI. If we studied people who are well educated logicians with little exposure to processing. Thus, those who are well trained in statistical judgment have developed a gist understanding of the meaning of “base rate.” A good logician does not need to process the verbatim details of each word in a syllogism. Rather, with a gist understanding of the bottom-line meaning of logical concepts such as the fallacy of the undistributed middle, *modus tollens*, and the like, he or she focuses on the key connectives such as “some” and “all” and ignores the rest. For unschooled participants such as ours, high scores on the logic index are associated with scores at the verbatim end of the FPPPI. If we studied people who are well educated logicians with little exposure to estimating probabilities we might make different predictions.

From the onset, the literature on the use of base rates has been characterized by controversies. For example, Kahneman and Tversky (1973) found small but statistically significant effects for base rates and concluded that, “prior probabilities were largely ignored when individuating information was made available” (p. 242). However, Ajzen (1977) conducted studies showing a somewhat larger base rate effect and concluded that, “in making predictions, people utilize information, including information supplied by population base rates, to the extent that they find it possible to incorporate the information within their intuitive theories of cause and effect” (p. 311–312). Medin and Edelson (1988) examined the use of experienced base rates in learning and categorization in diagnostic reasoning. They found that participants sometimes ignored base-rate, and in other circumstances used them appropriately or inappropriately (i.e. predicting that a rare disease is more likely than a common disease) depending on the category structure and test characteristics. Research by Reips and Waldmann (2008) further illuminated these results in a set of studies on predictive and diagnostic learning. They found that participants used base rate information with relatively simple research materials, but under conditions of greater complexity base rates were used only under conditions of diagnostic learning and not predictive learning. These controversies may be characterized as one of “base rate neglect” vs. “base rate respect.” On the one hand, there are decades of research and theory supporting the “base rate fallacy” (Bar-Hillel, 1980; Wolfe, 1995) or “base rate neglect” (Lyon & Slovic, 1976) indicating that people tend to ignore relevant base rates; and on the other hand, there is research and theory supporting “base rate respect” (Barbe y & Sloman, 2007) the position that people appropriately use relevant base rates at least under some circumstances when they are presented in an appropriate context or format. Our finding of reliable individual differences in the FPPPI may shed light on these contradictory findings in the literature. It may be that subtle task differences tap into fuzzy processing preferences in different ways. Further research using existing research materials and the FPPPI may illuminate apparent contradictions in the literature.

There are a number of other cognitive processes that could be fruitfully explored using the FPPPI. For example, taking an individual differences approach, using the FPPPI may shed light on Bayesian reasoning (Chater, Oaksford, Hahn, & Heit, 2010; Fisher & Wolfe, 2012) or even biases in reasoning and argumentation (Wolfe & Britt, 2008; Wolfe, Britt, & Butler, 2009). The FPPPI can also be used in applied research, particularly in medical decision making (Peters, McCaul, Stefanek, & Nelson, 2006; Reyna, 2008). We are interested in exploring how people differ in the extent to which they understand and respect or neglect the base rate information for breast cancer in making medical decisions such as whether to undergo testing for genetic breast cancer risk. We plan to use the FPPPI to assess the efficacy of interventions designed to improve statistical reasoning in a medical context (e.g. Wolfe, Hu, Reyna, & Fisher, 2011).

Finally, we have developed spreadsheets to calculate FPPPI scores using the items in Appendices A and B. We will make these spreadsheets available to researchers upon request.

### Acknowledgment

The project described was supported by Award Number R21CA149796 from the National Cancer Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health. The authors thank the National Cancer Institute for its support. The authors wish to thank the following for their competent assistance with data collection: Eric Cooke, Elizabeth Crawford, Emma Hogg, Talia Kenan, Dana Mitchell, Ben Nelligan, Kelly Papenfus, Ramy Toma, and Makinzie Utter.

### Appendix A. Nineteen fuzzy processing preference index items

| t parameter item means without base rates in parentheses |

**At Cloverdale High School** 10% of the seniors go on to college. Bob is a senior at Cloverdale High. He gets mostly As and Bs in school and is well liked by his teachers. What is the probability that Bob will go to college? (.891)

Researchers in Spain found that 20% of Spanish couples in their early 20s are still together after 3 months. Marco and Racquel, who are both 23 years old, recently initiated a relationship. Both like to go hiking on the weekends. They also have similar preferences in terms music and movies. What is the probability that they will be together after 3 months? (.685)

At Talia’s Restaurant the soup of the day is Cream of Cauliflower. The chef’s special recipe calls for plenty of Parmesan cheese. So far, 80% of the customers have described the soup as spicy. Brittany is just about to try the Cream of Cauliflower soup for the first time. What is the probability that Brittany will find the soup very spicy? (.278)

The Heights Tigers football team is playing their last game of the regular season. So far they have won 90% of their games. In their last practice they looked worse than they have all season. The players appear to be unmotivated and the coach sounds pessimistic. What is the probability that the Tigers will win their last game? (.395)

At Fisher Primary School, 15% of the 4th grade boys identify blue as their favorite color. Chris is a 4th grader at Fisher Primary School. He has a blue bicycle, and a blue helmet. What is the probability that Chris’ favorite color is blue? (.745)
At the Big Bird Supermarket 20% of the customers have a Big Bird Savings Card. Emma is a smart shopper who is often on the lookout for bargains. What is the probability that Emma has a Big Bird Savings Card? (.727)

At Jay’s Auto Sales, 90% of newly hired salespeople are able to meet their sales goal in the first week. During a recent interview at Jay’s Auto, Mark was described as nice, but unprepared and lacking in ambition. If Mark is hired, what is the probability that he will meet his weekly sales goal in the first week? (.429)

On the advanced deductive reasoning task, 10% answer at least 9 out of 10 questions correctly. Michelle enjoys solving puzzles and brainteasers. What is the probability that Michelle will correctly answer at least 9 out of 10 questions on the advanced deductive reasoning task? (.695)

Sam attended a speed dating service in which 10% of male participants receive the phone number of at least one female participant. Sam is considered good looking, likes new experiences, and is a natural leader in groups. What is the probability that Sam will receive at least one phone number? (.865)

20% of Jessop University female students spend an average of $100 a month on new clothing. Heather is a Jessop University student who follows fashion trends and has never been seen wearing the same outfit twice. What is the probability that she spends at least $100 a month on new clothing? (.825)

15% of Greek women get married before the age of twenty-five. Matina is a twenty-three years old Greek woman and has been in a serious, committed relationship for the past three years. What is the probability that she will be married by the time she turns twenty-five? (.746)

Frank is an outgoing teen and his family has a phone plan that allows unlimited data and phone calls. 20% of teens on his plan send at least 45 text messages per day. What is the probability that Frank sends at least 45 text messages per day? (.779)

A recent marriage survey found that 20% of newlyweds argue out of 10 questions correctly. Michelle enjoys solving puzzles and brainteasers. What is the probability that Michelle will correctly answer 8 out of 10 questions? (.30)

The town of Springfield has a nice botanical garden. What makes the garden unique is that it has dozens of wind chimes that make beautiful music in the soft summer breeze. About 70% of visitors to the garden report that they come to listen to the wind chimes. Mrs. Addison, who happens to be deaf, is visiting the botanical garden today. What is the probability that Mrs. Addison is visiting to listen to the wind chimes? (.70)

Jessica is a vegetarian. What is the probability that Jessica will order a hamburger? ____%

Since she was in first grade she has been on the soccer team and is already talking about colleges where she can get a soccer scholarship. Her favorite class is gym. What is the probability that Kate is overweight? (.197)

During the summer in New Haven, there is a 10% chance that it will rain on any given day. On July 1st in New Haven, it was mostly cloudy in the morning with dark clouds above. What is the probability that it will rain in New Haven at some point during that day? (.710)

WNL collects customer satisfaction information through a mail survey. Based on previous experience, 80% of customers return their surveys. Holly is a customer of WNL and is described by her friends as lazy, forgetful, and somewhat disorganized. What is the probability that Holly will complete and return the survey? (.275)

At South Gym, an anonymous survey indicated that 80% of the members have used steroids within the last year. Tony belongs to South Gym and cannot bench press 100 pounds. What is the probability that Toney has used steroids within the last year? (.214)

At the local bar and grill, wine and beer are served. According to the bartender’s records, 10% of all drinks sold are beer. Bill is a construction worker that describes himself as a “regular Joe.” Bill subscribes to Hotrod magazine. What is the probability that Bill will order a beer? (.779)

A rock band is conducting auditions for a new guitarist using a two-stage process. Based on previous auditions, 80% progress to the second stage. Andrew has been playing guitar for 2 months and has never performed even for friends and family. What is the probability that Andrew will progress to the second stage? (.296)

Appendix B. Four M-scale items

Richard is an avid skier and spends 90% of his vacations skiing. Today he has plane tickets to Aspen, Colorado and has been looking forward to this weekend trip for months. Unfortunately, Richard had a bad accident and both of his legs are broken. What is the probability that Richard will go skiing this weekend? ___

In Little Rock, Arkansas only 10% of the High School soccer referees are women. Sam has been a High School soccer referee for three years. Sam will not be refereeing this year because Sam is pregnant. What is the probability that Sam is a woman? ___

The town of Springfield has a nice botanical garden. What makes the garden unique is that it has dozens of wind chimes that make beautiful music in the soft summer breeze. About 70% of visitors to the garden report that they come to listen to the wind chimes. Mrs. Addison, who happens to be deaf, is visiting the botanical garden today. What is the probability that Mrs. Addison is visiting to listen to the wind chimes? ___

Jessica is on her lunch break. She doesn’t have much time so she decided to pop into McDonald’s for a quick lunch. About 80% of McDonald’s customers order some kind of hamburger. However, Jessica is a vegetarian. What is the probability that Jessica will order a hamburger? ___

References


