APPLIED COGNITIVE PSYCHOLOGY
Appl. Cognit. Psychol. **16**: 811–827 (2002)
Published online in Wiley InterScience 1 August 2002
(www.interscience.wiley.com) DOI: 10.1002/acp.830

Estimating National Populations: Cross-Cultural Differences and Availability Effects

NORMAN R. BROWN,[1]* XINJIE CUI[1] and ROBERT D. GORDON[2]

[1]*University of Alberta, Canada*
[2]*Brown University, Canada*

SUMMARY

Estimates of national population were studied in two experiments. In Experiment 1, Canadian and Chinese undergraduates rated their knowledge of 112 countries and then estimated the population of each. In Experiment 2, Canadians rated their knowledge of 52 countries and then provided population estimates for these *primed* countries and for a comparable set of 52 *unprimed* countries. In Experiment 1, participants from both nations produced estimates that resembled those obtained from Americans in prior studies (Brown and Siegler, 1992, 1993, 1996, 2001). However, there were several reliable cross-national differences in performance which appear to reflect cross-cultural differences in task-relevant naive domain knowledge. In addition, both experiments produced findings consistent with the claim that availability-based intuitions play an important role in this task. In Experiment 1, cross-national differences in rated knowledge predicted cross-national differences in estimated population; in Experiment 2, primed country names elicited larger population estimates than unprimed country names. We conclude by arguing for the general utility of this hybrid approach to real-world estimation. Copyright © 2002 John Wiley & Sons, Ltd.

It has been estimated that the Khmer Rouge killed over 1 million Cambodians between 1974 and 1978. As terrible as this figure is, one needs to know the population of the country at the time—approximately 6 million—to fully appreciate the enormity of the Pol Pot regime. In 1996, the Immigration and Naturalization Service estimated that some 2.5 million Mexicans were residing illegally in the United States. Does this figure indicate a mass exodus of Mexicans? Not really. After all, Mexico's population is about 100 million. Thus, the number of immigrants represents only a small fraction of the total. In 1998, the US government provided $67 million in aid to the Philippines and $61 million in aid to Nicaragua. These numbers might suggest that the US government is somewhat more concerned with the former country than with the latter. However, it is also true that 80 million people live in the Philippines, whereas only 4.7 million people live in Nicaragua. Thus, on a per capita basis, the USA was 16 times more generous to Nicaraguans.

The point of these examples is to demonstrate that accurate knowledge of national populations is essential for understanding world events, global trends, and government policy. Despite its importance, people, at least US university students, have little if any

*Correspondence to: Norman R. Brown, Department of Psychology, University of Alberta, Edmonton, Alberta T6G 2E9, Canada. E-mail: norman.brown@ualberta.ca

Contract/grant sponsor: NSERC.

explicit knowledge of national populations (Brown and Siegler, 1992, 1993, 1996, 2001). This fact raises several interesting questions. First, if population facts are unavailable, how do people estimate national populations? Second, is it only Americans who have a limited knowledge of national populations, or do people in other countries also lack this sort of demographic information? Third, assuming that Americans are not unique in their limited knowledge of national populations, do people in other countries use the same estimation strategies as Americans? Finally, if there are between-countries differences in the estimation accuracy, why do they occur?

In this article, we present two experiments designed to address these issues and to test specific theory-derived predictions. Experiment 1 compared population estimates collected in Canada and China; Experiment 2 looked at the effect of prior exposure to a country's name on the size of its estimated population. Together, these experiments were intended to provide converging evidence for the contention that *availability*-based intuitions can play a crucial role in real-world quantitative estimation (Brown and Siegler, 1992, 1993). Experiment 1 also allowed us to investigate the nature and origins of cross-national differences in estimation performance.

PRIOR RESEARCH ON POPULATION ESTIMATION

The present study extends prior research on population estimation (Brown and Siegler, 1992, 1993, 1996, 2001). This research has demonstrated that population estimates collected from US college students display a pronounced *availability bias*. This is a tendency to provide larger estimates for well-known countries than for less well-known countries with the same actual population. For example, both Iraq, a moderately well-known country, and Nepal, a fairly obscure country, had populations of about 18 million in the early 1990s. Despite their comparable sizes, the median estimate of Iraq's population (30 million) was three times greater than the median estimate of Nepal's population (10 million, Brown and Siegler, 1993, Experiment 2). More generally, this line of research has demonstrated that there is a moderately strong correlation between estimated population and rated country knowledge (our index of availability; also see Brown *et al.*, 1985), that estimated population correlates more highly with rated knowledge than it does with true population, and that the partial correlation between estimated population and rated knowledge is substantial even when true population and true land area are statistically controlled.

These findings have been taken to indicate that people rely on availability-based intuitions when they estimate national populations (Brown and Siegler, 1992, 1993). There are two key aspects to this position. The first is that people believe that better-known countries have larger populations than less well-known countries; and the second is that they apply this belief when estimating national populations, inferring that a country has a relatively large population when they know a lot about it and that it has a small population when they do not. After a country's relative population size has been determined, an estimate is produced by selecting a numerical value from the corresponding portion of the response range.

There are, of course, other ways to estimate national populations. These include direct retrieval and reconstruction. The former involves retrieving the target country's population directly from memory, and the latter involves the retrieving the population of related countries (or other relevant numerical facts such as the city populations) and using these values as numerical reference points to support plausible inferences. Both strategies should produce accurate unbiased estimates. However, as noted above, population estimates often are very biased. Moreover, people generally do not have access to the populations of most countries (Brown and Siegler, 1992, 1993, and see below). Thus, it seems unlikely that either direct retrieval or reconstruction are often used when people estimate populations. In contrast, numerical facts are not required to generate availability-based estimates and the nature of the observed biased suggests that availability-based intuitions do play a role in population estimation.

Note, we use the term availability here to maintain consistency with prior studies in this series (Brown and Siegler, 1992, 1993). There are, however, a number of mutually compatible processes capable of producing a subjective assessment of relative country knowledge. These include a process that gauges the ease with which country knowledge is brought to mind (this corresponds to one of Tversky and Kahneman's uses of the term (1973, 1974); cf. Betsch and Pohl, in press) and those that monitor familiarity or processing fluency. The present study was not designed to distinguish between these possibilities, but rather to provide converging evidence for the more general claim that population estimation is informed by a memory-based attribution process.

This latter claim is an important one. Not only is it intended to provide an accurate description of the way that people estimate national populations, it also implies that a cue-validity approach is necessary for understanding real-world estimation (Brown and Siegler, 1992, 1993). On this view, both domain-specific knowledge (e.g. facts about a country's land area, dominant religion, level of economic development) and heuristics (e.g. availability-based intuitions) are assumed to play a role in real-world estimation, though the former is not necessarily given more weight than the latter. Instead, estimates are hypothesized to reflect a weighted blend of competing sources of information, with the weighting of each source determined by its predictive strength (Brunswik, 1955). Thus, availability should play an especially important role when domain knowledge is sparse (as is the case in many laboratory demonstrations), and when it is more predictive of values on the target dimension than domain-specific knowledge. Conversely, availability should be given less weight than domain knowledge when the former is more predictive than the latter (see Friedman, 1996, for a similar analysis restricted to event dating). Brown and Siegler (1993, Experiment 1) tested these predictions in an experiment that compared population estimates (a task where availability is more predictive than other sources of domain knowledge) with estimates of national land area (a task where domain knowledge is more predictive than availability because knowledge of maps and globes provide valid domain specific information about areas; Kerst and Howard, 1978). As expected, participants displayed a pronounced availability bias when estimating populations, but not land area.

OVERVIEW OF THE PRESENT STUDY

In brief, the data consistent with the claim availability is assessed during the estimation process and used as an index of national population, and there are significant theoretical implications associated with this way of interpreting biased population estimates. Nonetheless, this position is subject to two types of challenges. First, it is possible that some factor related to but different than availability (e.g. GNP) underlies the observed correlations. Second, although we have assumed that people assess availability as part of the estimation process, the previous studies, which depended on correlational data, provide no direct evidence of this.

The present study addressed these problems. As noted above, Experiment 1 compared the population estimates and knowledge ratings produced by university students in China and Canada. Because prior research on this topic had been conducted in the United States, we were interested in determining whether people in other countries also relied on availability when they estimated national population. Moreover, China and Canada differ from one another culturally and geographically. This suggests that relative knowledge about various test countries might also differ (Pinheiro, 1998; Saarinen, 1989, 1999). In other words, we expected that participants' nationality would serve as a naturally occurring, within-item, availability manipulation.

If participants in both countries rely on availability when estimating populations and if there are marked between-nation differences in what is known about the test countries, then countries that are more familiar in China should elicit larger estimates from the Chinese than from the Canadians, and those that are more familiar in Canada should elicit larger estimates from the Canadians. It follows that cross-cultural differences in rated knowledge should correlate with cross-cultural differences in estimated population (see Brown and Siegler, 1992 for a parallel argument concerning the possible effect of longitudinal changes in rated knowledge on estimated population).

Experiment 1 was designed to capitalize on naturally occurring differences in availability. To provide converging evidence regarding the influence of availability on estimates of national population, Experiment 2 employed a laboratory-induced manipulation. In this experiment, test countries were divided into two similar sets. At the outset of the session, participants rated their knowledge of countries in one set (the primed set), but not the other (the unprimed set). Then, they estimated populations of all test countries. If the knowledge rating task increases the availability of the primed countries, and if people assess availability when they estimate national populations, then test countries should receive larger estimates when they are primed than when they are not.

Note, the logic behind the two experiments was much the same. In both cases, an attempt was made to vary availability within-items, but between-groups; and, in both cases, a key prediction was that population estimates for a given country would reflect between-group differences in availability, with countries eliciting larger estimates when they were better known (Experiment 1) or recently primed (Experiment 2). This (pseudo-) experimental approach was adopted because it is capable of demonstrating an availability effect in a way that separates item identity from differences in item knowledge (Brown and Siegler, 1992; Gabrjelcik and Fazio, 1984). As a result, data collected in these experiments should allow us to argue against the view that the biased population estimates observed in prior studies simply reflect some aspect of country knowledge that is correlated with, but different from, availability.

EXPERIMENT 1

In this experiment, Chinese and Canadian university students rated their knowledge of 112 countries and then estimated the current population of each. If participants from both countries rely on availability when making their estimates, results from this experiment should parallel those observed in prior studies carried out in the United States. Rated knowledge should strongly correlate with estimated population, and this relation should hold even when true population and land area are statistically controlled. In addition, for reasons outlined above, cross-national differences in rated knowledge should predict cross-national differences in estimated population (Brown and Siegler, 1992, 1993).

Data collected in this experiment provided a unique opportunity to compare estimation accuracy and rated knowledge between countries. In particular, we hoped to determine whether there were between-nation differences in estimation accuracy. We also were interested in possible differences in rated knowledge because these ratings might reveal nation-specific ‘world views’.[1] Our expectation was that these between-nations comparisons would produce explicable differences in both population estimates and knowledge ratings and that explanations developed to account for these differences might provide the basis for generalizations linking nationality, domain knowledge, and estimation performance. Because this was the first experiment to compare population estimates and knowledge ratings collected in different countries, it was necessary to treat this aspect of the study in an exploratory manner.

Experiment 1 did include a simple experimental manipulation. Prior to the estimation task, half of the Canadians were informed of the current population of Canada (27.4 million) and half were not. Likewise, half of the Chinese learned China’s current population (1165.8 million), and half did not. We did this because we were interested in determining whether exposure to a single *seed fact* would improve estimation accuracy (Brown and Siegler, 1993, 1996, 2001; Friedman and Brown, 2000a,b).

We expected that informed participants would produce more accurate estimates than uninformed participants, but that the difference would only be observed in measures that assess *metric* knowledge (i.e. knowledge of the statistical properties of the target dimension such as its mean, range, variance, and skew), and not in those that assess *mapping* knowledge (i.e. knowledge of relative size of the populations of different countries). This prediction was based on prior research demonstrating: (a) many real-world estimation tasks require both metric and mapping knowledge, (b) these two sources of knowledge are often independent, and (c) people react to *seed facts* by adjusting task-relevant beliefs to bring them in line with the information conveyed by seed facts (Brown and Siegler, 1993). In principle, a single *seed fact* could indicate whether one’s metric assumptions are incorrect, but carry no information about the relative country size. Thus, we expected this between-groups manipulation to produce differences in measures that reflect metric knowledge (i.e. order of magnitude error; see below), but not in measures that reflect mapping knowledge (i.e. the rank-order correlation between estimated and actual population).

Method

Material

This experiment was conducted in the summer and fall of 1993. At the time, the most recent *Information Please Almanac* (1993) indicated that there were 113 countries with 4 million or more people. With two exceptions, all these countries served as test items. Taiwan was excluded from the Chinese stimulus set because its independence was not recognized by the Peoples’ Republic of China. Czechoslovakia, which had recently

[1]In a pilot study Canadians and Americans rated their knowledge of 108 countries and estimated their populations. Two mean knowledge ratings were then computed for each stimulus country, one over the Canadian responses and a second over US responses. Then these means were correlated yielding an r of 0.98. Based on this result, we concluded that Canadians and Americans have a very similar world view.

dissolved, was excluded from the Canadian set because reliable population figures for the Czech Republic and Slovakia were not available when test booklets were being constructed. The 112 countries presented to the Canadians had a mean true population of 50.8 million, a median of 16.1 million and a skew of 6.47.

The 112 test countries were listed 56 to a page on the second and third pages of separate three-page knowledge-rating and population-estimation booklets; task instructions were printed on the first. A response field (a blank line) followed by the appropriate prompt ('knowledge' or 'millions') appeared next to each country name in each booklet. Random permutations were used to assign country name to position within different versions of the two booklets. Six Chinese versions of each booklet were created, and a separate English-language version of each booklet was created for each Canadian participant. Instructions were written in English and translated into Chinese.

Procedure

Participants first filled out the knowledge-rating booklet. They were instructed to work through the booklet, one country at a time, at their own pace, and to rate their current knowledge of each country on a *0* ('no knowledge') to *9* ('a great deal of knowledge') scale. The population-estimation task followed the knowledge rating task. Participants were instructed to provide their best estimate of the current population of each country and to take their best guess when they were unsure of a population. As before, participants were asked to work through the test countries, one at a time, in the order they were presented, and to take as much time as was necessary to provide accurate estimates. Half of the Chinese also received instructions that included information about the current population of China (1165.8 million), and half did not. Likewise, half of the Canadians were informed of the current population of Canada (27.4 million), and half were not.

Participants

Sixty Chinese and sixty Canadians participated in this experiment. The Chinese were undergraduates at Queuing University in Beijing; the Canadians were undergraduates at the University of Alberta. Advertisements placed around campus were used to recruit the Chinese participants, who were paid 10 yuan (approximately $2 Canadian) for their cooperation. Canadian participants, all of whom were born in Canada, were drawn from the introductory psychology subject pool and received course credit for their cooperation. Data were collected in small group sessions lasting about 45 minutes.

Results and discussion

Dependent measures

Unless otherwise noted, dependent variables were first computed within-subjects, over 109 test countries. Canada, China, Czechoslovakia, and Taiwan were excluded to maintain comparability between groups. Several accuracy measures are presented below. These include: proportion of correct estimates, median estimated population, signed Order of Magnitude Error (*OME*), absolute OME, and the rank-order correlation between estimated and actual population.

A response was scored as correct if it was within 5% of the test country's true population. The number of correct responses was then divided by the number of test countries to compute proportion correct. This measure provides an upper bound estimate for the presence of explicit, accurate population facts.

A signed OME was computed for each estimate according to the following formula: Signed OME = $\log_{10}$(estimated population/true population). An absolute OME was computed for each estimate by taking the absolute value of the signed OME. The OME measures convert estimation error to a percentage of an order of magnitude (Nickerson, 1981) and were used to assess the accuracy of participants' metric knowledge. The correlation between estimated and true population served as our index of mapping knowledge (see Brown and Siegler, 1992, 1993 for further discussion of these measures and their relation to the Metrics and Mappings framework). Rank-order correlations between estimated population and rated knowledge, estimated population and true land area, rated knowledge and true area, and rated knowledge and true population were also computed for each participant. Note, the correlation between rated knowledge and true population provided a measure of the validity of availability as a cue to national population.

All dependent measures were submitted to a Nationality (Canadian versus Chinese) × Instruction Type (national population presented in instructions vs no national population presented in instructions) Analyses of Variance.[2] Means for each dependent measure, computed over all Canadian participants and all Chinese participants, are presented in Table 1.

Overall accuracy

Population estimates collected from American students tend to be 'moderately good in a relative sense, but quite poor in an absolute sense' (Brown and Siegler, 1992, p. 409). The same can be said of those collected from Canadian and Chinese undergraduates. The data presented in Table 1 indicate that participants from both countries responded in a way that produced a moderate rank-order correlation between estimated and true population. At the same time, the dearth of accurate estimates (4%) indicates that participants, regardless of nationality, generally lacked specific knowledge of national populations, and the large OME values indicate that they also lacked accurate metric knowledge.

It appears then that the dissociation between quantitative and qualitative knowledge in this domain is not uniquely American. Rather, this pattern may reflect factors associated with national population *per se*. For example, people may have difficulty inducing the metric properties that characterize the distribution of national populations because they are rarely exposed to a representative sample of populations, and because specific population figures are hard to remember (Brown and Siegler, 1996). Also, these figures may be difficult to remember because they are rarely reported in the media, because they change constantly and, in some cases, quite rapidly, and because very large numbers are difficult to understand and hence difficult to encode in a meaningful way (Paulus, 1990).

Between-nation differences in accuracy

Although the general pattern of performance was similar in China and Canada, there were several reliable between-group differences. Interestingly, these differences did not indicate that participants in one country produced more accurate estimates than those in the other. Rather, they suggest that the Chinese had a better understanding of relative country size and the Canadians had a better understanding of the metric aspects of the task. Specifically, the rank-order correlation between estimated and true population was larger for the Chinese than for the Canadians (0.53 versus 0.43), $F(1, 116) = 21.79$, $MSE = 0.02$,

[2]ANOVAs were performed on *r*-to-*z*-transformed correlations. However, for the sake of clarity we present the back-transformed versions of the means computed for these analyses.

Table 1. Means (with standard errors) for Canadians and Chinese groups in Experiment 1

					Rank order correlations					Partial correlation
	Estimated population+	Signed OME**	Absolute OME**	Per cent correct	Est pop w/ knowledge***	Est pop w/ true pop***	Est pop w/ true area**	Knowledge /w true pop***	Knowledge w/ area**	Est pop w/ knowledge**
Canadians	20.84	−0.04	0.45	3.75	0.50	0.43	0.25	0.35	0.11	0.43
	(2.60)	(0.04)	(0.02)	(0.002)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)
Chinese	14.84	−0.24	0.54	4.22	0.60	0.53	0.31	0.45	0.17	0.50
	(1.79)	(0.05)	(0.03)	(0.003)	(0.03)	(0.02)	(0.02)	(0.01)	(0.01)	(0.03)
Mean	17.84	−0.14	0.49	3.98	0.56	0.48	0.27	0.40	0.14	0.46
	(1.59)	(0.03)	(0.02)	(0.002)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)

Note: *P*-levels for the difference between Canadian and Chinese means: $+p < 0.10$; $*p < 0.05$; $**p < 0.01$; $***p < 0.001$.

$p < 0.0001$. In contrast, absolute OME was smaller for the Canadians than the Chinese (0.46 versus 0.53), $F(1, 116) = 8.56$, $MSE = 0.03$, $p < 0.01$. Chinese also tended to produce smaller population estimates than Canadians (14.83 million versus 20.84 million); $F(1, 116) = 3.60$, $MSE = 299.76$, $p = 0.06$. The signed OME means make a similar point; on average, the Chinese underestimated populations by 0.24 of an order of magnitude and the Canadians by 0.04 of an order of magnitude; $F(1, 116) = 9.54$, $MSE = 0.12$, $p < 0.001$. Regardless of nationality, participants who produced more biased estimates, as measured by signed OME, also tended to produce less accurate estimates, as measured by absolute OME; absolute values of the mean signed OME was strongly correlated with the mean absolute OME for both Canadians ($r = 0.92$) and Chinese ($r = 0.89$).

Incidental, though explicable, between-nation differences in naïve domain knowledge may underlie both between-nation differences in the metric accuracy and the between-nation difference in correlation between estimated and true population. The between-nations difference in metric accuracy appears to be related to the fact that Canada (population 27.4 million, ranked 32nd) has a more typical population than China (population, 1165.8 million, ranked 1st). This suggests that knowledge of Canada's population might be more useful in this task than knowledge of China's.

More concretely, consider a Canadian who knew Canada's population, but no others, and who also assumed that this value fell somewhere in the middle of the range. This individual would be able to correctly infer that countries with small populations had fewer people than Canada, that mid-sized countries had populations similar to Canada's, and that large countries had considerably larger populations. Of course, knowledge of Canada's population and relative size do not provide precise information about the size of other populations. However, Canada's population would be useful as a mid-range reference point and would allow participants to partition the response range in a realistic manner. In contrast, knowing that China has the largest population in the world correctly implies that estimates for all other countries must be smaller than the value assigned to China, but leaves the further division of the response range unspecified, and hence subject to error. In the absence of additional information, there is simply no way of knowing whether a medium-sized country should be assigned a population of, say, 4 million, 40 million, or 400 million, as all of these figures are larger than 0 and less than 1 billion. As noted above, the Chinese tended to underestimate national populations. This suggests that these participants favored an ethnocentric division of the response scale, one that accentuated China's position as the world's most populous country.[3]

Seeding effects

In prior experiments (Brown and Siegler, 1993, 1996), exposure to specific, task-relevant numerical information (i.e. seed facts), increased the accuracy of metric beliefs and

[3]This explanation assumes that most participants had a good sense of the absolute and relative size of the population of their own country. There are data that support this assumption. First, the true population of China was presented to half the Chinese, and the true population of Canada to half the Canadians. Obviously, participants in these groups knew the exact populations of their own country before they began the estimation task. It turns out that most participants in uninformed groups also either knew their own population or could infer it accurately; 73% of uninformed Canadians and 70% of the uninformed Chinese responded to their own country with estimates that were within 10% of its true population. Finally, China elicited the largest population estimate from 59 of the 60 Chinese, and most Canadians provided estimates for Canada indicating that they understood that it has a mid-sized population. Specifically, across the 60 Canadians, the median rank for Canada's estimated population was 29.25.

decreased metric bias. The present study replicated these earlier findings and demonstrated that exposure to a single numerical fact—in this case, the population of one's own country—was sufficient to produce a reliable seeding effect. In both countries, participants who learned their country's true population prior to the estimation task displayed less bias and greater metric accuracy than those in the uninformed group; for signed OME, $F(1, 116) = 4.89$, $MSE = 0.12$, $p < 0.05$; for absolute OME, $F(1, 116) = 4.08$, $MSE = 0.03$, $p < 0.05$. Specifically, mean signed OME for informed and uninformed Chinese and informed and uninformed Canadians were −0.15, −0.32, 0.01, and −0.10, respectively. Absolute OME means were: 0.50, 0.58, 0.42, and 0.47. The Nationality × Instruction Type interaction was non-significant for signed and absolute OME ($F < 1.0$ in both cases). Follow-up comparisons indicated that there was a reliable effect of nationality on both measures, regardless of instruction type. Thus, although seeding effects were observed, the pattern of between-country differences was unaffected by exposure to the population facts.

With the exception of the findings just reported, none of the analyses discussed in this section or the next yielded a reliable effect of instruction type or a reliable Nationality × Instruction Type interaction. This is consistent with prior research that has demonstrated that seed facts may influence metric knowledge (i.e. OME measures) but not mapping knowledge (i.e. the correlation between estimated and true population). It also supports the notion that the two types of knowledge are independent (Brown and Siegler, 1993, 1996).

Availability effects

Both Canadians and Chinese produced population estimates that correlated more strongly with their knowledge ratings than with the true populations of the test countries. For the Canadians, the average rank order correlation between estimated population and rated knowledge and between estimated population and true population were 0.53 and 0.43 respectively; comparable figures for the Chinese were 0.60 and 0.50. A 2 (correlation type: estimated population with rated knowledge versus estimated population with true population) × 2 (nationality: Canadian versus Chinese) ANOVA confirmed that rated knowledge was a better predictor of estimated population than was true population, $F(1, 118) = 25.87$, $MSE = 0.02$, $p < 0.0001$, for correlation type. This analysis also indicated that both types of correlations were stronger in China than Canada, $F(1, 118) = 30.74$, $MSE = 0.04$, $p < 0.0001$, for nationality. The Nationality × Correlation Type interaction was not significant, $F < 1.0$.

Rated knowledge also correlated with true population (mean $r = 0.35$ and 0.45 for Canadians and Chinese, respectively, $F(1, 116) = 62.22$, $MSE = 0.01$, $p < 0.0001$; see below) and, to a lesser extent, with true land area (mean $r = 0.11$ and 0.16 for Canadians and Chinese, respectively, $F(1, 116) = 17.71$, $MSE = 0.004$, $p < 0.0001$). Because true population and land area covary with rated knowledge, a separate partial Spearman correlation was computed for each participant over the 109 test countries. These correlations assessed the strength of the relation between estimated population and rated knowledge, controlling for linear effects of true population and land area. The mean partial correlation was 0.50 for the Chinese and 0.42 for the Canadians, $F(1, 116) = 6.97$, $MSE = 0.04$, $p < 0.01$, with 57 of 60 Chinese correlations and 55 of 60 Canadian correlations being significant at the α-level of 0.05 or greater.

The final prediction regarding availability was that between-nation differences in rated knowledge would correlate with between-nation differences in estimated population. This prediction complements the preceding ones because it allows direct control over true population and true land area. To do this it was necessary to treat test country rather than participant as the unit of analysis. Thus, we computed a median population estimate and mean knowledge rating for each test country, separately for the Canadian responses and the Chinese responses.

Across the 109 test countries, the rank order correlation between the Canadian knowledge ratings and the Chinese knowledge ratings was 0.81 ($p < 0.0001$), as was the correlation between the Canadian and Chinese population estimates. The existence of this less-than perfect correlation established a necessary condition for determining whether knowledge differences predict differences in estimated population. To test this prediction, we computed two scores for each test country: a knowledge difference score (mean Chinese knowledge rating – Canadian knowledge rating) and an estimated population difference score (median Chinese population estimate – median Canadian population estimate). The rank-order correlation computed between these two difference scores over the 109 test countries was 0.52, $p < 0.0001$. Thus, above and beyond differences in metric beliefs, when a country was better known in China than Canada, it also tended to elicit larger population estimates in China (and vice versa), and when between-nation differences in rated knowledge were large, between nation-differences estimated population also tended to be large.

Between-nation differences in rated knowledge and their effects on population estimation

In previous sections, we noted that the correlation between estimated population and true population was stronger in China than in Canada and that the correlations between estimated population and rated knowledge and between rated knowledge and true population were also stronger in China. Given that rated knowledge is more predictive of true population in China than Canada, and assuming that the Chinese rely on availability as much as the Canadians, it follows that the Chinese should produce population estimates that correlate more strongly with true population than the Canadians. Moreover, from a cue-validity perspective, the fact that the Chinese knowledge ratings are better calibrated with true population than the Canadian ratings implies that the Chinese should rely on availability more heavily, and hence produce population estimates that are more strongly correlated with rated knowledge.

Of course, these explanations still leave open the question of why rated knowledge and true population correlate more strongly in China than Canada. To investigate this issue, we ranked each country twice, once according to its mean Chinese knowledge rating and once according to its mean Canadian knowledge rating. We then sorted the countries into three groups: a group of Asian countries ($n = 17$, excluding China), a group of European and English speaking countries ($n = 27$, excluding Canada), and the rest ($n = 65$). Paired t-tests indicated that the Asian countries generally were ranked higher (i.e. were better known) by the Chinese than the Canadians (23.56 versus 49.00; $t(16) = 7.66$, $SD = 11.7$, $p < 0.0001$), that the European countries were generally ranked higher by Canadians than Chinese (25.15 versus 36.04; $t(26) = 2.87$, $SD = 19.3$, $p < 0.01$), and that there was no systemic between-nation difference in the rankings of non-Asian, Third World countries (72.02 and 72.86 for the Canadians and Chinese respectively; $t(64) < 1.0$). Asian countries also tended to have larger populations than European countries; the mean true population for the countries in the two regions were 105.78 million and 37.00 million, respectively, with corresponding medians of 44.3 million and 10.5 million.

In combination, these facts explain why Chinese knowledge ratings correlated more closely with true population than Canadian knowledge ratings. Chinese gave relatively high ratings to large-population Asian countries and relatively low ratings to small-population European countries; the Canadians did the opposite.

More generally, the cross-national differences in rated knowledge support the notion that cultural similarity and physical proximity determined, at least in part, whether people are familiar with a country other than their own (Pinheiro, 1998; Saarinen, 1999). In addition, the rankings indicated that countries which have traditionally played a large role in world affairs (e.g. the United States was ranked 2*nd* by both Canadians and Chinese; Britain was ranked 4*th* by Canadians and 5*th* by Chinese; Japan was ranked 7*th* by Canadians and 4*th* by Chinese) and those that have been in the news for extended periods (e.g. Israel ranked 13*th* by Canadians and 19*th* by Chinese; Iraq ranked 16*th* by Canadians and 15*th* by Chinese) were well known in both Canada and China. We also found that countries from the non-Asian Third World (e.g. Paraguay was ranked 84*th* by Canadians and 77*th* by Chinese; Burkina Faso was ranked 109*th* by Canadians and 111*th* by Chinese) or had been formed following the (then) recent disintegration of the Soviet Union (e.g. Moldova was ranked 107*th* by Canadians and 103*rd* by Chinese; Tajikistan was ranked 103*rd* by Canadians and 93*rd* by Chinese) were obscure in both Canada and China.

EXPERIMENT 2

Experiment 1 demonstrated that people, regardless of nationality, often provide larger population estimates for well-known countries than less well-known countries of the same objective size. Moreover, as expected, between-nation differences in rated knowledge predicted between-nation differences in estimated population; when a country was better known in China than in Canada, it also tended to elicit larger estimates from the Chinese, and vice versa. This is an important result because it demonstrates that population estimates reflect naturally occurring fluctuations in availability, and because it unconfounds level of knowledge from the identities of the test countries.

These findings and those reported in prior studies (Brown and Siegler, 1992, 1993) are consistent with the notion that people consider availability and use it as an index of national population during the estimation process. Experiment 2, a priming study, was designed to test this claim more directly. Prior to this experiment, 104 test countries were assigned to one of two sets in a way that equated true population, estimated population, and rated knowledge across sets. The experiment itself consisted of three tasks and required two groups. During the first task, participants in one group rated their knowledge of all of the countries in one set, Set A, and none of the countries in the other, Set B; participants in a second group rated their knowledge of all Set B countries and no Set A countries. During the second task, all participants provided population estimates for both primed and unprimed countries. During the third task, they provided knowledge ratings for all of the countries in both sets.

In designing this study, we assumed that familiarity is reflected in population estimates and that the exposure provided by the initial knowledge rating tasks would increase familiarity for rated (i.e. primed) countries but not for unrated (i.e. unprimed) countries. It follows that participants should provide larger population estimates for primed countries than for unprimed countries. Thus, the average estimated population for a given test country should be larger when that country had appeared in the initial knowledge rating rated task than when it had not.

Method

Materials

When this experiment was conducted, in the spring of 1996, the most recent *Information Please Almanac* (1996) indicated that there were 116 countries with 4 million or more people. In a pilot study, 53 University of Alberta undergraduates provided knowledge ratings and population estimates for all these potential test countries. In order to produce two very similar stimulus sets, it was necessary to eliminate the following countries: China, India, the United States, Russia, Japan, Germany, Egypt, Canada, North Korea, Australia, Ecuador, and Malawi. The remaining 104 test countries were assigned to one of two sets, an A set and a B set. These sets were closely matched in terms of actual population (26.3 million versus 25.8 million), estimated population (15.6 million versus 15.4 million), and rated knowledge (1.9 versus 2.0); in all cases, $p > 0.10$.

Procedure

All participants performed three tasks. During the first task, Group A participants rated their knowledge of the 52 Set A countries, and Group B participants rated their knowledge of the 52 Set B countries. Countries were rated on a *0-to-9* scale, with *0* indicating 'no knowledge' and *9* indicating a 'great deal of knowledge'. Prior to the second task, participants were informed of Canada's current population (28.4 million) and were instructed to estimate, as accurately as possible, the current population of each test country to the nearest tenth of a million. As in Experiment 1, participants were warned that they might have difficulty with some of the questions and they were encouraged to take their best guess when they were unsure of an answer. The estimation task was followed by the second knowledge rating task, identical to the first knowledge rating except that participants were required to rate their knowledge of countries in the unprimed set as well as those in the primed set.

During each task, test countries were presented, one at a time, in unique random order. On a given trial, a country's name appeared in the centre of a computer controlled visual display along with the appropriate response field. To respond, the participant typed a number at his or her keyboard and then hit the ENTER key; this cleared the display, and initiated the next trial.

Participants

Sixty Canadian-born, University of Alberta undergraduates took part in this experiment. Half were randomly assigned to the *A Group*, and half to the *B Group*. Participants were tested individually in sessions lasting less than an hour and received course credit for their cooperation.

Results and discussion

As in Experiment 1, several measures were computed for each participant. These included the participant's median population estimate; mean signed OME; mean absolute OME; percentage correct; rank-order correlation between estimated and true population; rank-order correlation between estimated population and rated knowledge; and the partial correlation between estimated population and rated knowledge, controlling for the effects of true population and land area. Because all test countries were presented during the second knowledge rating task, only ratings collected during this task were used to compute the reported correlations. A set of preliminary *t*-tests indicated that there were no reliable between-group differences on any of these measures.

Overall performance

As in Experiment 1, the correlation between estimated and true population was moderate (mean $r = 0.36$), absolute OME was quite large (mean absolute OME = 0.53), and accurate responses were rare (percentage correct = 3%). Also as before, participants displayed a pronounced availability bias; estimated population correlated more strongly with rated knowledge (mean $r = 0.52$) than with true population (mean $r = 0.36$; $F(1, 59) = 38.82$, $MSE = 0.03$, $p < 0.0001$), and the partial correlation between estimated population and rated knowledge still remained substantial even when true population and land area were statistically controlled (mean $r = 0.46$).

Priming effect

The main aim of the present experiment was to determine whether there was a relation between recent exposure to country names and the magnitude of subsequent population estimates. Both subject-based and items-based analyses were carried out to assess this relation. To conduct the former, it was necessary to obtain each participant's median estimate for the 52 primed countries and the 52 unprimed countries. Similarly, for each test country, median estimates were computed for the 30 primed responses and for the 30 unprimed responses.

As predicted, primed estimates (mean = 23.3 million) were reliably larger than unprimed estimates (mean = 21.2 million; for subjects, $t(59) = 2.44$, $SE = 0.86$, $p < 0.02$; for countries; $t(103) = 1.89$, $SE = 0.52$, $p = 0.06$). This finding provides converging evidence for the claim that availability-based intuition plays a role in population estimation. It also demonstrates that even a single prior exposure is sufficient to produce a detectable difference in the magnitude of people's population estimates. Admittedly the observed effect was modest; primed estimates were only 9% larger than the unprimed estimates. Then again, the manipulation used in this experiment was quite subtle. Outside of the laboratory, exposure differences are far more extreme. For example, a recent full-text search of the 365-day New York Times Archive produced 2,684 hits for Israel (population 5.6 million) and 27 hits for Benin (population 6.1 million). It seems likely that these vast differences in exposure would produce large differences in familiarity, and that these differences, in turn, produce large differences in estimated population. In the present experiment, participants estimated that there were about 24 million people living in Israel and about 6 million people living in Benin.

GENERAL DISCUSSION

The present study was undertaken to address three related issues. First, because prior research on population estimation had been carried out in the USA, we were interested in determining whether Canadians and Chinese would produce the same general pattern of performance as Americans. Second, we hoped to obtain two sources of converging evidence for the claim that people rely on availability-based intuitions when estimating national populations. Finally, we wanted to determine whether Canadian estimates differed from Chinese estimates and to identify the origins of these differences, should they occur.

Data presented above addressed all three questions. Experiment 1 indicated that Canadians and Chinese performed much like Americans when they estimate national populations. Regardless of nationality, correlations between estimated and true populations tended to be moderately strong, absolute OME tended to be quite large, and accurate estimates were uncommon. Also, rated knowledge correlated more strongly with estimated population than with true population, and it predicted estimated population even when the influence of true population and land area were statistically controlled.

These data replicate findings from the United States and suggest that Chinese and Canadians, like Americans, rely on availability when estimating national populations. Moreover, between-nation differences in rated knowledge predicted between-nation differences in estimated population (Experiment 1), and countries elicited larger population estimates when they had recently been primed than when they had not (Experiment 2). Because both experiments were designed to demonstrate an availability effect in a way that unconfounds country identity from differences in country knowledge, these results rule out the possibility that biased population estimates are based on some aspect of country knowledge that is correlated with, but different from, availability. In addition, these results demonstrate that population estimates can be directly influenced by both naturally occurring and laboratory-induced differences in familiarity. It is worth noting that this study and its core predictions converge with others which have demonstrated that familiarity plays a role in a variety of tasks, including truth judgments (Begg *et al.*, 1992), fame judgments (Jacoby *et al.*, 1989), risk assessment (Lichtenstein *et al.*, 1978), and date estimation (Brown *et al.*, 1985; Kemp and Burt, 1998). Thus, the present research provides additional evidence that mere exposure can influence performance on a wide variety of real-world judgement tasks.

Although Canadian and Chinese population estimates were similar in many ways, we did find several reliable between-country differences. On the one hand, absolute OME was larger and signed OME more negative in China than Canada. On the other, the Chinese produced population estimates that correlated more strongly with the true population and rated knowledge than the Canadians. Also Chinese knowledge ratings correlated more strongly with true population than Canadian knowledge ratings. It appears that Canadians displayed better accuracy and less bias on the OME measures because Canada's population is more representative than China's, and hence more useful as a numerical reference point. It also appears that estimated population, true population, and rated knowledge were more strongly correlated in China because the Chinese were familiar with large-population Asian countries and unfamiliar with small-population European countries, whereas the opposite was true for the Canadians. These arguments led us to conclude that both discrepancies have their origins in incidental between-nation differences in naïve domain knowledge.

In the future, it will be interesting to collect population estimates from people living in many different countries. It would then be possible to determine whether we can identify significant cross-cultural differences in task-relevant knowledge and whether we can use these differences to predict when and how the data sets will differ from one another. For example, based on the present study, we predict that people who live in countries with very small populations (e.g. Micronesia, Suriname) or with very large populations (e.g. India, Russia) are likely to hold less accurate metric beliefs than those who live in countries with medium sized population (e.g. Argentina, Kenya). We can also predict that population estimates and knowledge ratings are likely to be better calibrated in Asia, where neighbouring countries tend to have relatively large populations, than in western Europe, where neighbouring countries tend to have relatively small populations.

Not only did the data collected in these two experiments allow us to address the issues that initially motivated this study, they provide support for several additional claims concerning population estimation and cross-national differences in world-view. First, the instructional effect observed in Experiment 1 demonstrated that exposure to a single numerical fact can improve the accuracy of metric beliefs and the difference. Second, the between-nation differences in metric accuracy (Canadians more accurate than Chinese) and mapping accuracy (Chinese more accurate than Canadians) provided additional evidence for the contention that these sources of knowledge are independent (Brown and Siegler, 1992; 1993). Third, the comparison of Canadian and Chinese knowledge ratings allowed us to identify the countries that were well known in both nations, that were obscure in both nations, and that were well known in one nation but not the other. Two types of countries fell into the first category: larger countries that have played and continue to play a major role in world affairs, and smaller countries that have been in the news for extended periods. With the other countries, knowledge ratings were related to cultural similarity and physical proximity (Pinheiro, 1998; Saarinen, 1988, 1999). Countries that were physically and culturally distant from both Canada and China received low knowledge ratings in both places, and those that were close to China but not Canada received high ratings only from the Chinese, and vice versa.

We would like to close with a methodological point. The present study combined experimental and exploratory methods and used both cross-cultural differences and laboratory manipulations to test specific parallel predictions. Admittedly, this was an unusual approach to take. But it turned out to be a highly informative one. We now have two additional sources of support for the availability hypothesis, and we now know that there is nothing uniquely American (or even North American) about inaccurate, biased, population estimates. Moreover, our attempt to understand the cross-national differences in population estimation yielded two general proposals: (a) with the exception of a small set of prominent countries, physical proximity and cultural similarity determine how much people of a given nationality will know about countries; (b) cross-national differences in estimation performance reflect cross-national differences in naïve domain knowledge. The first claim is consistent with recent findings in the human geography literature (Pinheiro, 1998; Saarinen, 1999). In contrast, to the best of our knowledge, the second claim is unprecedented, but does have testable implications. Given these contributions, we feel that the current project provides a nice demonstration that cross-cultural designs, in conjunction with more conventional experimental techniques, can be used profitably to investigate other real-world estimation tasks and that this hybrid approach should increase our understanding of how, why, and when cultural differences will be reflected in estimation performance.

ACKNOWLEDGEMENTS

This research was supported by an NSERC operating grant awarded to the first author. We thank Alinda Friedman, Peter Lee, and Bob Siegler for their comments on a draft of this article.

REFERENCES

Begg IM, Anas A, Farinacci S. 1992. Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General* **121**: 446–458.

Betsch T, Pohl D. in press. The availability approach to frequency judgment: A critical analysis. In *Frequency Processing and Cognition*, Sedlmeier P, Betsch T (eds). Oxford University Press: Oxford.

Brown NR, Rips LJ, Shevell SK. 1985. Subjective dates of natural events in very long-term memory. *Cognitive Psychology* **17**: 139–177.

Brown NR, Siegler RS. 1992. The role of availability in the estimation of national populations. *Memory & Cognition* **20**: 406–412.

Brown NR, Siegler RS. 1993. Metrics and mappings: a framework for understanding real-world quantitative estimation. *Psychological Review* **100**: 511–534.

Brown NR, Siegler RS. 1996. Long-term benefits of seeding the knowledge-base. *Psychonomic Bulletin & Review* **3**: 385–388.

Brown NR, Siegler RS. 2001. Seeds aren't anchors. *Memory & Cognition* **29**: 405–412.

Brunswik E. 1955. Representative design and probabilistic theory in functional psychology. *Psychological Review* **62**: 193–217.

Friedman A, Brown NR. 2000a. Reasoning about geography. *Journal of Experimental Psychology: General* **129**: 193–219.

Friedman A, Brown NR. 2000b. Updating geographical knowledge: Principles of coherence and inertia. *Journal of Experimental Psychology: Learning, Memory, & Cognition* **26**: 900–914.

Friedman WJ. 1996. Distance and location processes in memory for the times of past events. *The Psychology of Learning and Motivation* **35**: 1–41.

Gabrielcik A, Fazio RH. 1984. Priming and frequency estimation: A strict test of the availability heuristic. *Personality and Social Psychology Bulletin* **10**: 85–89.

Jacoby LL, Kelley CM, Brown J, Jasechko J. 1989. Becoming famous over night: limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology* **56**: 326–338.

Kemp S, Burt CDB. 1998. The force of events: Cross-modality matching the recency of news events. *Memory* **6**: 297–306.

Kerst SM, Howard JH Jr. 1978. Memory psychophysics for area and length. *Memory & Cognition* **6**: 327–335.

Lichtenstein S, Slovic P, Fischhoff B, Layman M, Combs B. 1978. Judged frequency of lethal events. *Journal of Experimental Psychology: Human Learning and Memory* **4**: 551–578.

Nickerson R. 1981. Motivated retrieval from archival memory. In *Nebraska Symposium on Motivation 1980*, Flowers JH (ed.). University of Nebraska Press: Lincoln; 73–119.

Paulus JA. 1990. *Innumeracy: Mathematical illiteracy and its consequences.* Vantage: New York.

Pinheiro JQ. 1998. Determinants of cognitive maps of the world as expressed in sketch maps. *Journal of Environmental Psychology* **18**: 321–339.

Saarinen TF. 1988. Centering of mental maps of the world. *National Geographic Research* **4**: 112–127.

Saarinen TF. 1999. The Eurocentric nature of mental maps of the world. *Research in Geographical Education* **1**: 136–178.

Tversky A, Kahneman D. 1973. Availability: a heuristic for judging frequency and probability. *Cognitive Psychology* **5**: 207–232.

Tversky A, Kahneman D. 1974. Judgment under uncertainty: Heuristics and biases. *Science* **185**: 1124–1131.

