The Bounds of Cognitive Heuristic Performance on the Geographic Profiling Task

PAUL J. TAYLOR1*, CRAIG BENNELL2 and BRENT SNOOK3

1Department of Psychology, Lancaster University, Lancaster, UK
2Department of Psychology, Carleton University, Ottawa, Ontario, Canada
3Department of Psychology, Memorial University of Newfoundland, St John’s, Newfoundland, Canada

SUMMARY

Human performance on the geographic profiling task—where the goal is to predict an offender’s home location from their crime locations—has been shown to equal that of complex actuarial methods when it is based on appropriate heuristics. However, this evidence is derived from comparisons of ‘X-marks-the-spot’ predictions, which ignore the fact that some algorithms provide a prioritization of the offender’s area of spatial activity. Using search area as a measure of performance, we examine the predictions of students (N = 200) and an actuarial method under three levels of information load and two levels of heuristic-environment fit. Results show that the actuarial method produces a smaller search area than a concentric search outward from students’ ‘X-marks-the-spot’ predictions, but that students are able to produce search areas that are smaller than those provided by the actuarial method. Students’ performance did not decrease under greater information load and was not improved by adding a descriptive qualifier to the taught heuristic.

Psychologists have long been interested in the cognitive mechanisms that allow people to provide effective solutions to complex, ill-defined problems. In recent years, attention has shifted to consider how the cognitive system handles problems derived from naturalistic settings such as criminal investigations (e.g., Bennell, 2005; Crego & Alison, 2004; Ormerod, Barrett, & Taylor, 2005). A problem that has attracted considerable attention is the geographic profiling task, where the goal is to predict an offender’s home location from knowledge of where his or her crimes were committed. A number of studies (Bennell, Snook, Taylor, Corey, & Keyton, 2007; Paulsen, 2006; Snook, Canter, & Bennell, 2002; Snook, Taylor, & Bennell, 2004; van de Kemp, van Ruth, Blokland, & Snook, 2005) have shown that human performance on this task matches the performance of computer algorithms, but that the process underlying this performance is based on a simple heuristic rather than a complex set of calculations. As found in other areas (Gigerenzer & Brighton, in press; Hope & Wright, 2007; Ormerod & Chronicle, 1999; Reich, Green, Brock, & Tetlock, 2007), human judgment on the geographic profiling task depends on a process that ignores much of the available information.

Copyright © 2008 John Wiley & Sons, Ltd.

*Correspondence to: Paul J. Taylor, Department of Psychology, Fylde College, Lancaster LA1 4YF, UK.
E-mail: p.j.taylor@lancaster.ac.uk
Recently, however, Rossmo (2005) has argued that human performance will fall short of algorithm performance when the full complexity of the geographic profiling task is considered. The concerns raised by Rossmo are examples of more general questions about the nature and limitations of cognitive heuristics. Briefly, they relate to (i) the capacity of heuristics to provide solutions to complex problems, (ii) the capacity of heuristics to cope with information load, and (iii) the capacity of heuristics to incorporate ‘exception’ rules into their definitions. Thus, the issues raised by Rossmo provide an agenda for exploring some of the capabilities and limitations of heuristic-led judgments. Consequently, they form the focus of this article. We begin by reviewing the geographic profiling task and the concerns raised by Rossmo, outlining how each reflects an unanswered question about the bounds of cognitive heuristics. We then report an experiment that evaluates the ability of people to make predictions under the scenarios suggested by Rossmo, while also allowing some exploration of the underlying heuristic rule(s) that drives performance on this task. As we will show, a direct test of individuals’ ability to interpret spatial behavior provides a further opportunity to understand the breadth and detail of the heuristics that individuals use to handle problems from naturalistic settings (e.g., Berntsen & Rubin, 2007; Chronicle, MacGregor, Ormerod, & Burr, 2006; Gigerenzer, Todd, & The ABC Research Group, 1999).

**Algorithms, heuristics, and the geographic profiling task**

The geographical profiling task requires the prediction of an offender’s home location from information on where his or her crimes were committed. The left panel of Figure 1 shows an example of the problem, together with the actual home location (‘H’) of the offender. The task is common in law enforcement settings where it may be used, among other things, to prioritize suspects, identify criminal strongholds, and focus door-to-door enquiries. The task is distinct from other decision-making tasks, such as the traveling salesman problem (Chronicle et al., 2006), because it requires an inference about the unknown from known information, rather than a direct evaluation of known information. It is also unique because there is no known computational method of finding the solution (i.e., making a perfect prediction about the location of an offender’s home). The task therefore falls into a class of

![Figure 1. Example of the geographic profiling task with the home location marked by an H (left panel) and a solution to this example as provided by a computerized actuarial method (right panel)](image-url)

increasingly studied tasks in which a correct solution cannot be guaranteed and cognitive shortcuts are required to circumvent the impossible processing demands of finding an optimal solution (Chronicle et al., 2006; Smith & Gilhooly, 2006).

Perhaps because of the intractable nature of the geographic profiling task, the standard solution since the early 1990s has been computational. Researchers and practitioners have used dedicated computer packages (e.g., Rigel; Rossmo, 2000) to apply an algorithm that derives ‘likelihood-of-home’ values for areas around each of the offender’s crimes. By summing these likelihoods across the crime series, the algorithm produces a probability surface on which a higher value indicates a greater predicted likelihood of finding the offender’s home at a given location. The right panel of Figure 1 presents an example of the outcome of the actuarial process for the map presented in the left panel of Figure 1. The grid cells shown on this map are each assigned a likelihood value. These likelihoods are made easier to interpret by concentric bands of color. Each band represents different likelihoods of locating the offender (e.g., top 10%, 10–20%, etc.), with the highest band indicated by the darker shade in the center of the map (colored red in the original output). The letter ‘L’ marks the point identified by the actuarial method as the area that is most likely to contain the offender’s home (i.e., the point of highest likelihood).

In an effort to better understand how people solve this particular problem, and in response to evidence of the efficacy of cognitive heuristics (Gigerenzer et al., 1999), several recent studies have compared the predictions made by the algorithms to predictions made by students and police officers. Against many expectations, these studies have shown that adults using a simple heuristic can make predictions that are comparable, and sometimes better than, those made by the algorithms used in actuarial tools (Bennell, Snook, et al., 2007; Paulsen, 2006; Snook et al., 2002; Snook et al., 2004). For example, the results of Snook et al. (2004) suggest that many untrained adults are able to make satisfactory predictions based on a basic understanding of one of two patterns of criminal spatial activity. These spatial regularities are summarized by a decay heuristic (i.e., offenders commit offences close to home) and a circle heuristic (i.e., offenders’ homes are located in a circle whose diameter is the distance between the furthermost crimes). After one of these two heuristics is taught to participants, almost all make predictions that are as accurate as the actuarial method. It appears that people are able to interpret the visual pattern of crime locations in a way that makes efficient use of information without having to attempt large and cumbersome calculations.

**Geographic profiling as a prioritization problem**

According to Rossmo (2005), there are several reasons why it may be premature to claim that heuristics provide an equally accurate and more frugal solution to the geographic profiling task. One reason to be cautious, he argues, relates to the likelihood surface that is produced by an actuarial approach. This surface provides an investigator with a ‘search strategy’ (Rossmo, 2000), which is an explicit map of the ‘optimal’ way to search an area to find the offender’s home. Specifically, because the actuarial method predicts the likelihood

---

1Interestingly, while each of these heuristics has been shown to produce accurate predictions, the prediction that they produce on any individual map can be very different. For example, when presented with a map where the crimes are clustered to one side of the activity space, a participant using the circle heuristic will predict towards the center of all the locations, while a participant using the decay heuristic will predict away from the center and toward the cluster of crimes. Indeed, in this scenario, a participant using the decay heuristic may make a prediction near to the cluster of crimes that falls outside of the circle drawn around the furthestmost crimes.
of an offender living at each location in the activity space, it is possible to search systematically through the space by considering locations in the order of decreasing likelihood (Canter, Coffey, Huntley, & Misen, 2000). Thus, the effectiveness of the actuarial method relates not to the prediction of a single point (i.e., an ‘X-marks-the-spot’), but to the amount of area that is searched before locating the offender’s home. This measure of effectiveness is quite different from that employed in previous studies of human and actuarial performance. In a typical geographic profiling task, a participant marks an ‘X’ at the location they believe is the offender’s home, and this is compared to the point of maximum likelihood as predicted by the algorithm (i.e., the ‘L’ in Figure 1). Since participants do not provide a search strategy, the performance of the algorithm and participants is measured as a straight-line distance between their prediction and the offender’s home location; a measure known as error distance.

While the importance of search strategy will be small in cases where the strategy mirrors a uniform search away from the point of highest likelihood (i.e., an expanding circle), it may be quite different in other cases. For example, consider a crime series that is characterized by a period of offending close to home and a second period of offending that is in a location far from home. This crime series will result in a bimodal distribution, where crime locations cluster in two separate areas. A heuristic-led judgment in this scenario would likely place a prediction midway between these two clusters of crimes and, consequently, a significant distance away from the actual home location. In contrast, an actuarial prediction provides a search strategy that focuses equally on points around each of the two crime clusters. The result is that the home may be in the location of second-highest likelihood, but that this point appears on the opposite cluster to the point of highest likelihood. The area searched in this scenario will be very small, but the error distance will (inappropriately) be much larger.

It is this kind of possible scenario that has led several researchers to argue that existing error distance comparisons unfairly mask the effectiveness of the actuarial method (Rich & Shively, 2004; Rossmo, 2005). This possibility has led these researchers to argue that an appropriate comparison must consider the amount of area that is searched before locating the offender’s home (i.e., search area\(^2\)), and not simply the distance between the actual home location and a single ‘best’ prediction of the home location. In the absence of an equivalent human prediction, this comparison would involve searching outward in a uniform manner from a participant’s ‘X’ prediction (similar to ripples from a pebble dropped in water), identifying the smallest circle that incorporates the offender’s home, and comparing the area of this circle to the area searched when following the actuarial search strategy. Accordingly, when this approach is taken, these researchers believe that the actuarial method will provide a more effective solution to the geographic profiling task than human judges.

Based on the argument that actuarial methods would outperform human judges, we predict that:

**H1:** Participants introduced to the Circle or Decay heuristic will perform significantly worse than an actuarial method when performance is measured by a naïve search area (i.e., the smallest fitting circle).
UNCOVERING THE BOUNDS OF GEOGRAPHIC PROFILING HEURISTICS

Incorporating search area into studies of geographic profiling is a potentially important step forward in efforts to understand performance on this task. However, inherent in a comparison of actuarial methods and human predictions is the presumption that people are limited to predicting a single location of maximum likelihood. In the same way that measuring accuracy with error distance masks the performance of actuarial methods, so comparing ‘X’ predictions to search strategies masks the possibility that people are able to use heuristics to provide an accurate ‘search area’. Thus, support for the hypothesis presented above (H1) may not necessarily indicate the superiority of the actuarial method over simple heuristics. Instead, it forces us to question the limits and character of cognitive heuristics.

The predictive bounds of cognitive heuristics

The notion that actuarially derived search strategies will outperform human predictions is treated largely as self-evident and obvious by advocates of computational systems (see Gigerenzer & Brighton, in press, for examples of this ‘more-is-better’ ideology in other areas). However, studies of prediction tasks in other areas suggest that individuals may possess some ability to prioritize options, and that perhaps this ability may transfer to providing search strategies on the geographic profiling task. Specifically, a range of research suggests that individuals are able to prioritize information by comparing each available option to a perceived or evidence-based ideal (Kee et al., 2002; Lee, Chandrasena, & Navarro, 2002; Taylor, Johnson, & Allen, 2004). For example, Lee et al. (2002) demonstrated that individuals have developed ‘one reason’ strategies for searching through PsycINFO outputs, and that this type of search often outperforms an exhaustive search of the outputs. Similarly, research on economic and social decisions, both in real-world and laboratory-based tasks, has demonstrated that individuals are able to prioritize options based on limited consumption of the available information (Todd & Gigerenzer, 2003).

So how might this type of cognition facilitate an accurate analysis of the distributions of crime locations? Given the perceptual nature of the geographic profiling task, it seems reasonable to suspect that the heuristics that facilitate good performance (e.g., the circle heuristic) are rooted in the way that humans perceive shapes. In our previous work, we have speculated that this may involve adaptive application of a gestalt-like judgment of the ‘center of gravity’ of a set of points (Snook et al., 2004). This is consistent with computational work that suggests shape identification is best achieved through judgments of the distance between key points on the diameter of the perceived shape, and the perceived center of that shape (Zhang & Lu, 2001, 2002). In the same way, a rapid perception of the center of crime locations is likely to form the basis of individuals’ ‘X’ predictions, particularly after this strategy is reinforced by heuristic training. Assuming this perception relies on judgments of the relative locations of points, it would then be possible to systematically extend the original judgment to indicate the direction in which the search should focus. Such an extension would be based on a simple interpretation of the ‘pattern’ of the relative locations of points, in much the same way that individuals interpret the pattern of locations when solving the traveling salesman problem (Chronicle et al., 2006). As long as this interpretation leads individuals to orient searches towards the majority of crimes, then they will provide search strategies that are effective. Given this possibility, we hypothesize:
**H2:** Participants will be able to provide a search area that predicts the home location of an offender as accurately as the search area provided by an actuarial method.

**Information bounds of cognitive heuristics**

A second concern presented by Rossmo (2005) relates to the number of crime locations that are used when making a prediction. Rossmo argues that the complex algorithms that are used in geographic profiling systems are most effective when there is information on more than five crime locations. Specifically, he suggests that ‘the more points (i.e. crime locations), the more information, and the greater the precision…. A minimum of five crime locations are necessary for stable pattern detection and an acceptable level of investigative focus’ (Rossmo, 2005, p. 652). The implication of this argument is that the actuarial approach will significantly outperform the heuristic approach when the necessary amount of information is available. Theoretically, we view this argument as addressing the issue of whether heuristic performance improves with increasing crime information, or whether, in contrast, heuristic performance is limited and does not make use of (or indeed gets worse as a result of) more information. Given the limited capacity of the human cognitive system (Broadbent, 1957), it is conceivable that additional information may increase the complexity of the geographic profiling task to a point where the individual’s cognitive processing becomes overloaded (Eppler & Mengis, 2004; Jacoby, Speller, & Berning, 1974). This is certainly a weakness alleged of cognitive heuristics in other areas, and it is the argument indirectly levied at heuristic performance on the geographic profiling task (Rossmo, 2005). Empirically, the question suggests that a more appropriate comparison of algorithm and heuristic performance would include a larger number of crimes than the three crimes examined in the Snook et al. (2004) experiment.

The idea that increasing cognitive load will reduce heuristic performance relative to actuarial performance is arguably at odds with the simple heuristic explanation of human performance on the geographic profiling task. As Gigerenzer and Brighton (in press) note, the feature that distinguishes heuristic solutions from computational solutions to the task is that they are calculated on the basis of only two of the crime locations (the two furthest from each other). This feature remains true regardless of the number of crime locations, such that large numbers of locations will impact on performance only in the sense that it will take more time to identify the two relevant locations on the map. This is only likely to occur when the number of crime locations is significantly large (in the traveling salesman problem significantly large is greater than 60; Chronicle et al., 2006). Consistent with this reasoning, Bennell, Snook, et al. (2007) found no significant relationship between the number of crimes presented to participants and the accuracy of their heuristic-led judgments. Specifically, using error distance as the measure of performance, they found the accuracy of participants’ predictions to be largely equivalent when based on three and seven crime locations. In light of this finding, our prediction is that:

**H3:** The accuracy of the search area provided by participants will not be significantly affected by the presentation of more than five crime locations.

**Definitional bounds of cognitive heuristics**

The final concern raised by Rossmo (2005) relates to the extent to which the heuristics provided to participants correspond to what is known about the regularities in offender
spatial behavior. Rossmo questions the extent to which the heuristics that were taught in previous research match with what is observed in criminal spatial behavior (i.e., the heuristic’s ecological fit). Specifically, Rossmo proposes that previously taught heuristics only approximate the ecological reality of offender spatial behavior and that modifying the heuristics to better match the pattern of offender behavior will improve human performance. To achieve this, Rossmo suggests that each of the previously taught heuristics will better represent theoretical knowledge if they include a qualifier about their application. Rossmo offered his own definitions of the Circle heuristic and the Decay heuristic. His revised Circle heuristic is ‘Half or more of criminals live within the circle encompassing their crime locations’, and his revised Decay heuristic is ‘Many offenders live close but not too close to their crime scenes’.

The definitions provided by Rossmo (2005) are interesting in several respects. First, they represent specific descriptions of the ecological reality of offenders’ spatial behavior, which contrast the general rules of thumb that arguably characterize the previous heuristics. The extent to which heuristics must tightly match the behavioral reality, versus simply reflect the general pattern of behavior, has not been widely considered, although there is some suggestion that a greater match between environment and heuristic improves the effectiveness of performance (Gigerenzer et al., 1999; Todd, Fiddick, & Krauss, 2000). Second, to achieve their ecological fit, the heuristics proposed by Rossmo contain a qualifier of the original rule. For example, the proposed Circle heuristic takes the original notion of offenders living within a circle drawn around the crime locations and makes the qualification that this fact is not true for all offenders (i.e., ‘half or more of criminals’). This qualification is empirically true, since some offenders (e.g., burglars) are found to live outside the circle (i.e., commuters, Canter & Larkin, 1993). Yet, as the studies by Berry and her colleagues illustrate (Berry, 2006; Berry, Knapp, & Raynor, 2002; Natter & Berry, 2005), individuals are often very bad at operationalizing qualifiers in a way that matches their original meaning (i.e., is ecologically rational). Thus, while improving the ecological fit of the heuristics should lead to better predictions, it remains interesting to explore the extent to which individuals can successfully apply heuristics that incorporate different levels of detail and complexity.

In an attempt to test the ecological fit of the heuristics proposed by Rossmo (2005), we predict:

**H4:** The accuracy of the search strategy provided by participants will be significantly improved by adding qualifiers to the taught heuristic rule.

**The current study**

The current study seeks to further understand the bounds of cognitive heuristic performance on the geographic profiling task. To achieve this, we return to the geographic tasks studied in previous research and address Rossmo’s (2005) concerns about the validity of the findings from this work. Using a new, larger sample of responses, we first assess the importance of search strategy in the task by evaluating whether or not the strategies provided by an actuarial method can outperform a non-strategic search outward from participants’ ‘X’ predictions. We then determine whether or not participants are also able to provide their own accurate search areas by comparing their proposed prioritization of locations to the search strategy derived from the actuarial method. In this comparison, performance is measured as the amount of area searched before locating the offender’s
home. To provide an exhaustive test under this measure, we compare the performance of human and actuarial methods under different information loads and under various definitions of the heuristics.

METHOD

Participants

Participants were 200 undergraduate students recruited on a voluntary basis from The University of Liverpool, UK. All were aged over 18 years and were allocated to one of the following five conditions: Control ($n = 40$), STB-Circle ($n = 40$), STB-Decay ($n = 40$), R-Circle ($n = 40$), and R-Decay ($n = 40$). As described below, the conditions prefaced by ‘STB’ received training in a heuristic proposed by Snook et al. (2004), while the conditions prefaced by ‘R’ received training in one of the alternative heuristics proposed by Rossmo (2005). None of the participants reported ever having police employment or experience with geographic profiling.

Materials

In line with previous studies, we recorded participants’ predictions using an experimental booklet. The first page of the booklet was a general instruction sheet that described the geographic profiling task and asked participants to complete the task by marking an ‘X’ at the point where they felt the offender was most likely to live. These instructions were followed by a second ‘heuristic advice’ page, which provided the experimental manipulation. Table 1 summarizes the advice given on this page for each of the five experimental conditions. For four of the groups, the sheet presented a sentence describing a heuristic, either the circle or decay heuristics used in Snook et al. (2004), or the circle or decay heuristic proposed by Rossmo (2005). The fifth group was given a blank page and we intended to use the performance of this group as a baseline.

Table 1. Experimental manipulation and origin for the five heuristic conditions

<table>
<thead>
<tr>
<th>Name</th>
<th>Origin</th>
<th>Rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>No information given</td>
<td></td>
</tr>
<tr>
<td>STB-Circle</td>
<td>Snook et al. (2004)</td>
<td>The majority of offenders’ homes can be located within a circle with its diameter defined by the distance between the offender’s two furthermost crimes</td>
</tr>
<tr>
<td>STB-Decay</td>
<td>Snook et al. (2004)</td>
<td>The majority of offenders commit offences close to home</td>
</tr>
<tr>
<td>R-Circle</td>
<td>Rossmo (2005)</td>
<td>Half or more of criminals live within the circle encompassing their crime locations</td>
</tr>
<tr>
<td>R-Decay</td>
<td>Rossmo (2005)</td>
<td>Many offenders live close, but not too close, to their crime locations</td>
</tr>
</tbody>
</table>

3Bennell, Snook, et al. (2007) report that undergraduate student performance on the geographic profiling task is comparable to the performance of police officers.
The booklet then presented 28 maps showing the crime site locations of different serial offenders. These maps were a combination of the stimuli presented in Snook et al. (2004) and Bennell, Snook, et al. (2007). Specifically, 10 of the maps presented the location of the first three victims of solved serial murder cases in Germany (see Snook et al., 2004). The remaining 18 maps presented the location of solved serial burglaries committed in St John’s, Newfoundland, Canada (see Bennell, Snook, et al., 2007). Of these maps, six presented three crimes, six presented five crimes, and six presented seven crimes. These maps were presented because they provided a direct re-examination of the conditions that prompted Rossmo’s concerns. This allowed us to consider the bounds of heuristics on the geographic profiling task while ruling out the possibility that any disparity between current and previous findings might be due to uncontrolled variation in stimuli.

To remain consistent with the information used by the computational technique, the maps were presented in black and white and without any topographical features. The maps were presented in a random order, but this order was kept consistent for all participants. Finally, to ensure that map orientation had no effect on performance, we rotated the maps by 180° for half (n = 100) of the participants (no differences were found).

The last page of the booklet invited participants to provide their own search area. The page instructed participants to revisit the 28 maps and make a prediction about the direction in which they would focus their search if they were the police. Specifically, the instructions explained the distinction between searching uniformly away from a predicted ‘X’ point (hereafter ‘naïve search’) and searching strategically (non-uniformly) away from a predicted point (hereafter ‘strategic search’). This page of the booklet then instructed participants to provide a strategic search area by drawing a closed line (irregular closed shape) around each of their original predictions. This line was to be drawn in such a way that it indicated the relative direction in which a participant felt it best to search outwards from their original prediction. For any one point of the drawn shape, a greater distance away from the original prediction, relative to the equivalent distances for other points, would indicate that participants felt an emphasis should be given to searching in that direction, relative to the other directions. A shorter distance away from the original prediction relative to the other distances would indicate a de-emphasis of that search direction. The instructions indicated that the size of the suggested search area was not important, only the relative directions that were emphasized for the search. The instructions encouraged participants to ask questions if they were unsure about what was required in the second stage of the experiment.

Procedure

Participants were tested individually at a table in a small room. They were presented with the experimental booklet and asked to read the instruction sheet until they understood the experimental task. They were invited to ask questions and reminded that they were free to leave the experiment at any time. After confirming that they understood the instructions, participants were asked to work their way through the booklet at their own pace. Once the participant had completed the booklet, and before they started working through the booklet for the second time, the experimenter reiterated verbally the strategic search instructions. This was to ensure that participants fully understood the concept of drawing a shape that emphasized their preferred search direction. Once the participants indicated that they understood the new instructions, they were asked to go back through the maps and make a strategic search prediction. After completing all 28 maps, the participants were thanked for...
their participation and debriefed as to the purpose of the study. Completion of both tasks took approximately 25 minutes.

**Obtaining computerized predictions**

To be consistent with previous research, we derived computer-based predictions for the 28 maps using a negative exponential function (Canter et al., 2000). This function, whose formula is presented elsewhere (Snook et al., 2004), assumes that the probability of locating an offender’s home residence decreases with increasing distance from his or her offences. As well as being a widely used function for geographic profiling, it has also been shown to have comparable performance to more elaborate functions (e.g., truncated negative exponential function, see Paulsen, 2006; Snook et al., 2005; Taylor, Bennell, & Snook, 2002).

We implemented the function by inputting the \(x\) and \(y\) coordinates of each crime location into the spatial statistics program CrimeStat (Levine & Associates, 2000). Consistent with other programs, CrimeStat superimposes a grid over the total map area and computes for each cell the straight-line distance between the cell and the crime locations. The derived distances are each submitted to a negative exponential function and the resulting values summed to provide a single value for each cell. These summed values are then transformed into a probability or likelihood value for the cells. The values indicate the predicted likelihood of finding the home location in the cell being considered relative to other cells in the grid. CrimeStat thus derives a likelihood value for each cell in the grid, allowing a researcher to determine both the cell with the highest likelihood (i.e., ‘X-marks-the-spot’ prediction) and a strategy for searching through cells of decreasing likelihood (i.e., strategic search). In the current study, CrimeStat implemented 7000 grid cells that covered an area of 235 mm \(\times\) 163 mm, which is the size of the standardized maps provided to participants.

**Measuring error distance**

To examine the impact of the different heuristic training on participants’ predictions, as well as to determine whether or not the data were comparable with observations from previous studies, we compared the accuracy of ‘X-marks-the-spot’ predictions using error distance. This was measured in millimeters as the straight-line distance between the actual home location and the predicted home location. For the actuarial method, the predicted location was taken as the cell assigned the highest likelihood by CrimeStat. For the participants, the predicted location was taken as their ‘X’ as marked in the experimental booklet. A larger value in this error distance indicates a less accurate prediction of the offender’s home location.

**Measuring search area**

We measured search area in millimeters as the amount of area that would be searched before locating the offender’s home. For the actuarial method, we counted the number of cells whose assigned likelihood was higher than the likelihood assigned to the cell containing the actual home location. We then divided this number by the total number of cells (i.e., 7000), and multiplied this proportion by the total activity space (i.e., 38 305 mm\(^2\)) to obtain a measure of the area that would be searched prior to finding the home location.

To obtain a naïve search area for participants (i.e., when they do not have an opportunity to draw a strategic area), we took our measure of error distance as the radius of a circle and
applied the standard geometric formula for calculating the area of a circle. This calculation is based on the notion that a non-strategic search of an offender’s activity space would radiate outward from the predicted point in a uniform (circular) manner. That is, we assume the non-strategic search strategy acts as ripples radiating out from the point at which a pebble is dropped in water, where the pebble reflects a participant’s original ‘X’ prediction. Thus, by using error distance as the radius, we obtain the area that would need to be searched prior to the offender’s home location appearing on the diameter of the circular search area.

To measure the area that would be searched when following participants’ strategic preferences, it was necessary to expand or shrink the drawn shape so that it incorporated the home location. Specifically, the drawn shape needed to be scaled so that the circumference of the shape passes through the offender’s home location; this being the minimum area searched before finding the home location. We achieved this scaling and derived a measure of search area for a particular map in three stages. First, we derived the size of participants’ proposed search areas, using graphical paper to aid our measurement of the area within the irregular shape. Second, we divided the participant’s error distance by the radius between the participant’s first prediction and the point on the circumference of their predicted area that was in line with the home location. This provided a scaling ratio (either above or below 1.00) that indicated the extent to which it was necessary to expand or reduce the drawn shape in order that its circumference passes through the offender’s home location. Third, the square of this ratio was multiplied by the area of the participants’ drawn search area to determine the area that would have been searched using the participants’ proposed search shape.

Design

We examined the performance of participants and the actuarial method using three dependent measures: error distance, naïve search area, and strategic search area. For each measure we made two types of comparisons. The first was a comparison of participant performance, which we achieved using a 5 (Heuristic condition: Control × STB-STB-Circle × STB-Decay × R-Circle × R-Decay) by 3 (Number of crimes: Three × Five × Seven) mixed design. The second was a comparison of participant performance in each condition to the performance of CrimeStat. Consistent with previous research (Bennell, Snook, et al., 2007; Snook et al., 2004), we achieved this using a series of one-sample t-tests in which participants’ performance across the maps was compared to the mean accuracy of the CrimeStat algorithm.

RESULTS

Error distance comparisons

Figure 2 shows the mean error distances for predictions made after training for the five heuristics. A 5 (Heuristic condition) by 3 (Number of crimes) mixed ANOVA revealed a main effect of Heuristic condition, $F(4, 195) = 6.72, \eta^2 = .14, p < .05$; a main effect of Number of crimes, $F(2, 390) = 348.09, \eta^2 = .77, p < .05$; but no significant interaction between Heuristic condition and Number of crimes, $F = 1.51$, n.s. To determine whether or not the heuristic training improved the accuracy of predictions, as found in previous
studies, we planned comparisons of the Control group performance to the performance of each of the heuristic groups. Compared to the Control group ($M = 46.1$, $SD = 13.9$), there was significantly better performance by the STB-Decay group ($M = 40.9$, $SD = 11.1$, $p < .01$), a non-significant trend of better performance for the STB-Circle group ($M = 42.5$, $SD = 11.6$, Dunnett’s $t$, $p = .12$), but no differences in the performance of the R-Circle ($M = 50.4$, $SD = 18.0$) and R-Decay ($M = 47.9$, $SD = 14.6$) groups (Dunnett’s $t$, $p > .96$).

Figure 2 also shows the mean error distance for the predictions of CrimeStat’s negative exponential function ($M = 43.7$ mm). The performance of CrimeStat was significantly better than the R-Circle group, $t(119) = 4.10$, $p < .01$, and the R-Decay group, $t(119) = 3.15$, $p < .01$, and marginally better than performance in the Control group, $t(119) = 1.92$, $p = .057$. There was no significant difference in the performance of CrimeStat and the STB-Circle group, $t(119) = -1.16$, n.s., but significantly worse performance by CrimeStat compared to the STB-Decay group, $t(119) = -2.78$, $p < .01$.

Naïve search area comparisons (H1)

To test the difference between actuarial and heuristic predictions when performance is measured using search area, we first compared the area searched when using the actuarial strategy to the area searched when moving naively away from participants’ ‘X’ predictions (i.e., concentric circles radiating out from the predicted ‘X’). Figure 3 shows the average search area for participants’ naïve search area predictions and actuarial predictions as a function of Heuristic condition and Number of crimes. As can be seen in Figure 3, the relative performance of the five heuristic conditions matches the pattern of performance shown in Figure 2, since the area of the naïve circle is proportional to the square of participants’ error distances (i.e., radius of the circle). Consistent with this, a 5 (Heuristic condition) by 3 (Number of crimes) mixed ANOVA confirmed comparable main effects of
Heuristic condition, $F(4, 195) = 6.04$, $\eta^2 = .12$, $p < .05$, and Number of crimes, $F(2, 390) = 175.72$, $\eta^2 = .42$, $p < .05$, but no significant interaction between Heuristic condition and Number of crimes, $F < 1$, n.s.

Consistent with H1, the actuarial strategies led to significantly less area being searched compared to the participants’ na"ıve search strategies. When compared to the area searched when using the strategy from CrimeStat ($M = 5726.8$), there was significantly more area searched when using the predictions from the Control group ($M = 9606.8$, SD = 5885.0), $t(119) = 7.22$, $p < .01$, the STB-Circle group ($M = 11861.0$, SD = 4339.2), $t(119) = 5.81$, $p < .01$, the STB-Decay group ($M = 7428.3$, SD = 3998.0), $t(119) = 4.66$, $p < .01$, the R-Circle group ($M = 11861.0$, SD = 9317.1), $t(119) = 7.21$, $p < .01$, and the R-Decay group ($M = 10366.1$, SD = 6512.9), $t(119) = 7.80$, $p < .01$.

Strategic search area comparisons (H2)

Two participants, both exposed to the STB-Decay heuristic, failed to mark a search area on one of their maps (Map 14 and Map 15, respectively). We nevertheless retained their data for the remaining 27 maps. The average search area for participant and actuarial predictions as a function of Heuristic condition and Number of crimes is shown in Figure 4. A 5 (Heuristic condition) by 3 (Number of crimes) mixed ANOVA revealed a main effect of Heuristic condition, $F(4, 195) = 9.19$, $\eta^2 = .19$, $p < .01$; a main effect of Number of crimes, $F(2, 390) = 42.09$, $\eta^2 = .42$, $p < .01$; but no significant interaction between Heuristic condition and Number of crimes, $F < 1.54$, n.s. Planned comparisons revealed that, compared to the Control group ($M = 6643.4$, SD = 5500.9), there were non-significant trends of better performance for the STB-Circle group ($M = 4792.5$, SD = 4031.0, Dunnett’s $t, p = .16$), and the STB-Decay group ($M = 4938.3$, SD = 3493.3, $p = .20$). There were no differences in performance between the Control group and the R-Circle group.
The planned comparisons of participants’ strategic search areas and the actuarial strategies revealed mixed support for the predicted equivalence in predictive accuracy (H2). When compared to the area searched when using the strategy from CrimeStat ($M = 5726.8$), there was marginally more area searched when using the predictions from the Control group ($M = 6643.4$, $SD = 5500.9$), $t(119) = 1.83$, $p = .07$, and significantly more area searched when using the predictions of the R-Circle group ($M = 7539.2$, $SD = 8001.0$), $t(119) = 2.48$, $p < .05$, and the R-Decay group ($M = 10962.6$, $SD = 10261.1$), $t(119) = 5.59$, $p < .01$. However, in contrast, there was significantly less area search when using the predictions of the STB-Circle group ($M = 4792.5$, $SD = 4031.0$), $t(119) = -2.54$, $p < .05$, and the STB-Decay group ($M = 4938.3$, $SD = 3493.3$), $t(119) = -2.47$, $p < .05$.

**Comparisons across information load (H3)**

Figures 2–4 also present the mean performance for predictions made on maps containing 3, 5, and 7 crime locations. As can be seen from these Figures, predictions were significantly better when based on five crime locations than when based on three or seven crime locations, regardless of whether performance was measured by error distance (Figure 2) or search area (Figures 3 and 4). The performance of participant and actuarial methods was the least accurate when based on three crime locations, more accurate when based on seven crime locations, and the most accurate when based on five crime locations. This ordering of performance is inconsistent with the inverse relationship expected between the number of crime locations and predictive accuracy.

---

4Under a two-tailed analysis, the Rossmo decay heuristic (R-Decay) performed significantly worse than the Control group (Dunnett’s $t$, $p < .001$).
presented crimes and the relative accuracy of human predictions (H3). To test Rossmo’s (2005) contention that actuarial methods can only perform effectively when based on more than five crime locations, and the corollary that human performance will worsen when presented with over five crimes, we ran ANOVA contrasts to compare performance on three and five crime locations with performance on seven crime locations. This contrast revealed no difference in performance for any of the participant groups or the actuarial method when performance was measured by error distance (all $t(117) < 1.2, \text{n.s.}$) and strategic search area (all $t(117) < 1.3, \text{n.s.}$).

Comparisons across new and old heuristics (H4)

To examine the relative performance of the original Snook et al. (2004) heuristics, and the modified heuristics proposed by Rossmo (2005), we compared accuracy across the equivalent old and new heuristic conditions (e.g., STB-Circle compared to R-Circle). Planned $t$-tests revealed that the STB-Circle group performed significantly better than the R-Circle group when performance was measured by error distance, $t(238) = -4.07, p < .01$, naive search area, $t(238) = -4.09, p < .01$, and search area, $t(238) = -3.36, p < .01$. Equivalent tests revealed that the STB-Decay condition performed significantly better than the R-Decay condition when performance was measured by error distance, $t(238) = -4.19, p < .01$, naive search area, $t(238) = -4.21, p < .01$, and strategic search area, $t(238) = -6.09, p < .01$.

DISCUSSION

In response to evidence suggesting that computational algorithms perform no better than human judges using heuristics, Rossmo (2005) raised a number of questions about the conditions under which such evidence is valid. His questions relate to the bounds of heuristic performance in the geographic profiling context (Simon, 1956), and they have both theoretical and practical implications. Theoretically, the questions concern whether or not heuristics are able to facilitate judgments that move beyond ‘X-marks-the-spot’ predictions to more sophisticated assessments of the crime locations. In one sense, this type of prediction task is far removed from the forced-choice tests and other cognitive tasks on which heuristics are known to perform well (Gigerenzer et al., 1999). Yet, it is a form of prediction (i.e., prioritization based on incomplete information) that is likely to be encountered in everyday life, and thus it is of theoretical interest to determine whether or not individuals have available heuristics that produce accurate solutions to this form of cognitive task. Practically, the question concerns whether or not there is a role for computational techniques in the activities of law enforcement officers and crime analysts. Use of computational geographic profiling is often expensive (Snook et al., 2004) and, in an environment where the demand for financial resources is fierce, it is important to ensure that there is a genuine need to augment police expertise with computer-based support. These issues motivated the current study to consider three of the questions raised by Rossmo (2005).

The unimportance of search strategy

The first question we addressed centers on the fact that actuarial methods provide a strategic way of searching an offender’s activity space. This aspect of actuarial solutions,
which has been overlooked in previous research, may significantly reduce the amount of area searched for an offender’s home. This question opens up the possibility that complex computational methods will perform better than human judges when search strategy is taken into account, raising doubts about the validity of previous findings. Consistent with this possibility, we found that an actuarial search strategy led on average to less area being searched to find an offender’s home compared to a uniform search outwards from participants’ ‘X’ predictions. This finding arguably lends support to those who believe it is important to consider the amount of area searched when measuring performance (Rossmo, 2005).

The trouble with this initial comparison is that it tips the balance of fairness back in favor of actuarial methods, since it assumes that participants are unable to provide a search strategy that is any more sophisticated than an ‘X-marks-the-spot’ prediction. However, as we argue in the Introduction, no evidence to our knowledge suggests that this is likely to be the case, and some indirect evidence suggests that heuristics might allow judges to provide an effective search area (e.g., Lee, Loughlin, & Lundberg, 2002). To test this possibility, we gave participants the opportunity to propose their own search area and compared the size of that area, to the size of the area produced by the actuarial tool. In excess of our expectations (H2), participants taught the STB-Circle or STB-decay heuristics provided a prioritization of the activity space that often outperformed the search strategies provided by the actuarial method.

How are participants achieving this level of accuracy? While it is difficult to provide a definitive answer to this question, our findings do allow us to draw some tentative conclusions. These conclusions stem from the observation that participants made their search area predictions at a pace similar to that at which they made their initial ‘X’ predictions. This suggests that both sets of predictions are based on a frugal evaluation of the crime sites and not on any detailed calculation of relative distances. Moreover, the fact that the STB-Circle and STB-Decay groups outperformed the Control group when making their search area predictions, as well as their initial ‘X’ predictions, highlights the importance of heuristics to performance on these tasks. Their importance to both tasks illustrates the adaptive nature of heuristics, which is consistent with research on other heuristics and other judgment and decision-making tasks (Martignon & Schmitt, 1999). It also raises the possibility that performance on the two tasks relies on the same underlying process. The process that seems common to both tasks is a scan for the crime locations that are furthest apart from one another, combined perhaps with an estimate of their midpoint (Zhang & Lu, 2001, 2002) and knowledge about the equity (i.e., minimize travel effort) that drives everyday spatial decisions (Jankowski, Andrienko, & Andrienko, 2001). An evaluation of performance using eye-tracking methodology will be one useful way of unpacking the processes that underlie application of the heuristics.

The effects of more information

As well as testing whether or not participants could add a search area to their ‘X’ predictions, we also considered the robustness of their predictions under high, medium, and low information load. In contrast to our hypothesis, we observed an ‘inverted-U’ relationship between the number of presented crime locations and the accuracy of student and actuarial predictions. For both the students and actuarial method, predictions were found to be most accurate when based on five crimes, less accurate when based on seven crime locations, and least accurate when based on three crimes. The only exception to this
pattern was actuarial performance when measured by search area, which was worst when based on seven crime locations rather than three locations. Despite this exception, the relationship between information and performance was strikingly uniform across the heuristic and CrimeStat conditions. This similarity suggests that information load plays an important role in performance, while the similarity in prediction patterns suggests that both approaches are relying on the same known regularities in offenders’ spatial behavior, albeit exploiting them in different ways. This serves to indirectly confirm that participants’ heuristics are providing a frugal way of making a prediction that is equivalent to the computational output of CrimeStat.

While the change in performance across number of crimes indicates that information load plays an important role, it does not support Rossmo’s (2005) proposal that actuarial performance will improve when predictions are based on greater numbers of crimes. Nor does it support the related implication that heuristic predictions will worsen when greater number of crimes are presented. Instead, the current findings suggest that there is an optimal amount of information on which to base predictions, and that predictions made outside of this ideal may suffer from either too much or too little information. While this conclusion is consistent with research on information overload (e.g., Eppler & Mengis, 2004; Jacoby et al., 1974), it should be considered tentative. It does not, for example, incorporate findings relating to the performance of participants on maps containing large numbers of crimes (i.e., more than seven crimes). Similarly, it does not rule out the possibility that there is something unique about the relationship between where the offender lived and where they committed their crimes on each of the maps containing five crimes (see Bennell, Snook, et al., 2007). Finally, it does not consider the possibility that we distorted our measure of participants’ performance on maps showing three crimes by averaging their performance across maps of burglary and murder locations. At the aggregate level, this possibility is largely ruled out by the fact that our findings across three-, five-, and seven-crime locations closely match those of Bennell, Snook, et al. (2007), whose participants only viewed maps of burglary locations. However, in light of recent evidence that murderers and burglars travel from their home base in different ways (Goodwill & Alison, 2005), we cannot dismiss the possibility that participants are applying the heuristics in subtly different ways across crime types, and that these individual differences are leading them to be more accurate on one type of crime compared to the other. Thus, while overall performance matches that observed in previous research, individual performances may be showing considerable variation across crime types. The sensitivity of heuristics to such differences in the pattern of locations, and the possibility that people may change their heuristic use to match these patterns, is an important area for future research.

The bounds of heuristic definitions

An important concern of research into cognitive heuristics is the extent to which the performance of a heuristic is determined by its ‘ecological fit’ with what is observed in the environment (Todd, Fiddick, & Krauss, 2000; Todd & Gigerenzer, 2003). Our final

Interestingly, however, there was little difference across the maps in the proportion of home locations occurring inside rather than outside of a circle drawn around the outermost crimes. Specifically, 56% were found within the circle on maps containing three crimes, 50% were found within the circle on maps containing five crimes, and 50% were found within the circle on maps containing seven crimes. Thus, there is a need to determine whether some other aspect of crime distribution improves heuristic performance or whether, as tentatively found in this study, the geographic profiling task is best performed with limited information.
hypothesis considered this issue by examining the relative performance of the heuristics used in previous research and the revised heuristics proposed by Rossmo (2005). These revised heuristics contained qualifiers of the original heuristics, which potentially gave them a better ecological fit to known patterns in offender spatial behavior. However, in contrast to the hypothesis, we found that these revised heuristics did not improve performance on the geographic profiling task, and that in several scenarios they led individuals to provide significantly worse predictions than those not taught the heuristic (i.e., the R-Decay group compared to the Control group).

These findings provide strong evidence to suggest that Rossmo’s (2005) revised heuristics should not be used as an alternative to those proposed by Snook et al. (2004). However, it is not clear why the revised heuristics did not lead to better performance. We suspect that the revised heuristics are not better representations of offenders’ spatial behavior. This might be true, for example, in the case of the Decay heuristics, since a number of recent studies have shown no difference in accuracy between algorithms that do and do not incorporate a ‘not too close’ element (Paulsen, 2006; Taylor et al., 2002). A second more attractive explanation may lie in individuals’ difficulties with interpreting the content of the rule itself. There is a good deal of evidence to show that people vary considerably in their interpretation of qualifiers when they are presented verbally rather than numerically (Berry, 2006; Berry et al., 2002). We suspect that participants’ interpreted the original heuristic term ‘majority’ as universal in meaning and so applied the heuristic to all maps. In contrast, participants may have chosen to make more qualified interpretations of the statements ‘half or more’ or ‘not too close’ and this led them to incorrectly apply the heuristic. It will be interesting to see whether more accurate predictions emerge when the refined heuristics are presented in a different form (e.g., numerically) or when participants are made to actively process the meaning of the heuristic (Natter & Berry, 2005).

CONCLUSIONS

While search area may be a more valid measure for assessing the effectiveness of actuarial methods of geographic profiling, it continues to be the case that people can predict the home location of an offender with a level of accuracy that is similar to that achieved by a computationally expensive actuarial method. Humans, it seems, are able to adapt and extend a simple rule about criminal spatial behavior to make judgments about the spatial distribution of a set of crime locations. Specifically, not only are they able to provide an ‘X-marks-the-spot’ prediction, but they are also able to predict a search area that, on average, incorporates an offender’s home more efficiently than the search strategy provided by a computer algorithm.

In line with previous research, our findings suggest that police agencies with limited access to geographic profiling software (either because of limited funding or because of time constraints) may be able to meet their requirements by training officers to use simple prediction rules (Petrocelli, 2007). Officers who receive this training may then apply these rules to a set of crime locations in much the same way as the participants did in the current study. In operationalizing this approach, it will be important for police agencies to consider a number of issues. One important issue is whether they should take steps to ensure that the officers making the predictions do not (consciously or unconsciously) apply a heuristic in a way that is biased towards an existing line of enquiry. To avoid this possibility, it may be prudent to ensure that predictions are made by an officer who has no involvement in the
investigation, using a map without topographic information so that local areas are not identifiable. As a further stop-guard against this possibility, it may also be useful to train officers in how to apply a simple actuarial technique (e.g., applying the circle heuristic using a protractor) (Bennell, Taylor, & Snook, 2007). Such techniques do not require computer software to deploy, and they are less susceptible to the variability in performance that may occur with human judgments.

Of course, these suggestions apply equally to professional geographic profilers, who may themselves be able to produce accurate reports in support of an investigation without the use of a computational system. Indeed, professional profilers may bring to bear an ability to fine-tune their prediction based on other offence information that sets their performance apart from our participants (cf. Ekman, O’Sullivan, & Frank, 1999). They may also provide advice on other related problems not dealt with in our geographic profiling task, such as when to stop enlarging the proposed search area (in the case of heuristic led predictions) or at what likelihood to stop searching through the cells (in the case of a computational output). This kind of fine-tuning may be valuable to a police agency, although its capacity to improve accuracy beyond simple heuristics has yet to be demonstrated. Examining the predictions of professional profilers will establish their effectiveness beyond the heuristic predictions examined here. It will also add to the growing literature that is mapping out the bounds of performance on complex, real-world problems.

ACKNOWLEDGEMENTS

Support for the research reported in this article was provided to the second author by the Office of the Dean, Faculty of Arts and Social Sciences, Carleton University, and to the third author by the Natural Sciences and Engineering Research Council of Canada.

REFERENCES


Bounds of geographic profiling heuristics


