Law Hum Behav (2008) 32:526–535
DOI 10.1007/s10979-008-9128-x

ORIGINAL ARTICLE

How Variations in Distance Affect Eyewitness Reports and Identification Accuracy

R. C. L. Lindsay · Carolyn Semmler · Nathan Weber · Neil Brewer · Marilyn R. Lindsay

Published online: 6 February 2008

Abstract Witnesses observe crimes at various distances and the courts have to interpret their testimony given the likely quality of witnesses' views of events. We examined how accurately witnesses judged the distance between themselves and a target person, and how distance affected description accuracy, choosing behavior, and identification test accuracy. Over 1,300 participants were approached during normal daily activities, and asked to observe a target person at one of a number of possible distances. Under a Perception, Immediate Memory, or Delayed Memory condition, witnesses provided a brief description of the target, estimated the distance to the target, and then examined a 6-person target-present or target-absent lineup to see if they could identify the target. Errors in distance judgments were often substantial. Description accuracy was mediocre and did not vary systematically with distance. Identification choosing rates were not affected by distance, but decision accuracy declined with distance. Contrary to previous research, a 15-m viewing distance was not critical for discriminating accurate from inaccurate decisions.

Keywords Eyewitness identification · Lineup · Description · Distance

R. C. L. Lindsay (✉) · M. R. Lindsay
Queen's University, 62 Arch Street, Kingston, ON, Canada K7L 3N6
e-mail: rod.lindsay@queensu.ca

C. Semmler
University of Adelaide, Adelaide, Australia

N. Weber · N. Brewer
Flinders University, Adelaide, Australia

In Illinois v Levin (1991) an eyewitness observed a stranger walking around the side of a neighbor's house. Later police determined that the house had been broken into and burgled, and Jeffrey Levin was arrested based on the description of the man and his car provided by the eyewitness (despite the fact that the detailed description of the car contained some clear discrepancies when compared with Levin's car). The identification of Levin by the witness was an integral part of the case. The witness estimated that he was approximately 25 m from the criminal when he observed him. Because the witness was looking out of a specific window and the criminal was seen walking on a walkway, it was possible later to measure the actual distance. The witness had been approximately 67 m from the criminal. The actual distance between the witness and criminal often will be determined by asking the witness to estimate how far he or she was from the criminal. In some cases it will not be possible to determine objectively if the witness' report is accurate. This raises several interesting issues that are examined in this study: How can the courts know what the actual distance was from the witness to the culprit in cases where objective markers may not exist? How can the courts determine the degree to which they should rely on identification evidence as distance or reported distance varies? Are actual distance, perceived distance, and remembered distance equally predictive of identification accuracy?

Interest in perception of distance has a long history. Ptolemy used the principle of clutter (filled distances are seen as larger than empty distances) in an attempt to explain the moon illusion, and experimental research on the phenomenon of distance perception started at least as early as the middle of the nineteenth century (Thorndyke 1981). A century later, researchers were directly studying the psychophysics of distance perception (e.g., Harway

1963; Kunnapas 1960). Wiest and Bell (1985) reviewed 70 studies of distance perception and memory (61 published, including 17 remembered distance) and concluded that both perceived and remembered distance conformed reasonably well to Steven's power law (Stevens 1957). Witmer and Kline (1998) state that "a common finding is that people are not very accurate at estimating distances, even to familiar objects (p. 145)." Perception of distance frequently produces underestimates (Witmer and Kline 1998; Wright 1995). Wiest and Bell reported an average 15% underestimate of distance across 45 studies. This literature suggests that witnesses may frequently underestimate the actual distance between themselves and criminals, although the estimate relied upon in the criminal setting is remembered rather than perceived distance.

The previous findings were based on perception of distance. Research on memory for distance suggests that remembered distance estimates are less accurate than perceived judgments of distance (Wiest and Bell 1985). One explanation for this effect is that the distance is "re-perceived" when the information is retrieved from memory using the same principles as distance perception (e.g., Kerst and Howard 1978). The claim is that the stored information represents an underestimation and when re-perceived generates a further (perhaps 15%) underestimate of actual distance.

The first issue examined in the present study was the accuracy of witnesses' judgments of the distance of a person whom they observed briefly. People were recruited in public places and asked to observe target persons from a variety of distances. Distance judgments were made immediately after viewing or from memory up to 48 h (maximum 48 h; minimum 30 min, $M = 12.6$ h, $SD = 9.18$ h) later. Second, we examined another issue previously unexplored in eyewitness research, namely, how the reliability of witness descriptions of the target person varied with viewing distance.

The third, and the major, issue examined was the variation of identification performance with distance. The effect of distance on processing of faces has been clearly demonstrated by Loftus and Harley (2005) who showed that as a face is moved away from an observer the available face details become progressively coarser. Thus, it may seem obvious that distance between a witness and the criminal should be an important determinant of the likely accuracy of a subsequent identification attempt. Perhaps this explains the paucity of research directly relevant to this issue. However, distance need not reduce accuracy of identifications per se. As distance increases, witnesses may realize that they were unable to accurately encode the appearance of the criminal and simply decline to choose from lineups (i.e., offer a 'not sure' response). This could result in a reduction in the frequency of identification without reducing the probability that a choice is correct (as a 'not sure' response should be considered neither correct when the target is absent nor incorrect when the target is present). Choosing behavior from lineups as distance increases is thus an important topic that has received virtually no attention.

Some research suggests that people are capable of recognizing familiar faces from a substantial distance. For example, Greene and Fraser (2002) showed participants photos of celebrities at 200 feet and then brought the photos closer by steps of 20 feet until the face was recognized (data for faces never recognized were excluded from the analyses). Participants recognized the faces of celebrities from about 160 feet (49 m) for women and just under 200 feet (62 m) for men. Such data suggests that successful facial recognition is possible at quite large distances (although, see Loftus and Harley 2005). Of course, eyewitnesses often are required to select unfamiliar faces from lineups. Wagenaar and van der Schrier (1996) exposed participants to faces at different distances (3–40 m) and, immediately after, asked them to attempt to select the target from a 6-person lineup. Recognition performance declined as distance increased. The authors concluded that performance was optimal (defined as producing a ratio of 5 or more hits per false alarm) only if the viewing distance was less than 15 m. Wagenaar and van der Schrier (1996) suggested a 15-m rule as a useful "rule of thumb" for the courts. A replication using photos of famous people led to the same conclusion (De Jong et al. 2005).

However, many questions remain unanswered. The Wagenaar studies examined the effect of actual distance on facial recognition accuracy. If actual distance is underestimated by witnesses, then courts may err if they rely on the 15-m rule (and witness reports of distance) as witnesses reporting that they were 15 m from the criminal may have been much further away. Further, no distance estimation study appears to have used a human as the "target," and thus it is possible that distance estimation from one person to another is more or less accurate than the distance estimation from a person to an inanimate object. Another possibility is that the witness may infer distance from the strength of the memory of the target. If this is so, the relation between remembered distance and identification accuracy may be the same as the relation between perceived distance and identification accuracy even if remembered distance is less accurate than perceived distance. Finally, the notion of a rule, even a rule of thumb, may suggest to the courts that an exclusionary principle be adopted, that is, identification evidence should be admitted only if the viewing distance does not exceed 15 m. Such an approach suggests that identification accuracy is generally good at distances up to 15 m and uniformly bad beyond 15 m, which seems unlikely. To examine these issues, our participant witnesses also were asked to attempt to identify the target person from a lineup.

Method

Participants

Participants were approached if they appeared to be adults but otherwise there were no restrictions. A total of 1,321 participants were recruited for the study. Of those who provided the information (eight unspecified), 54% of the final sample was female. Participants ranged in age from 17 to 82 ($M = 27.77$, $SD = 13.08$). Given the large sample size, it was possible to perform analyses based on age. The final sample included mostly people of European ancestry (82.4%); however, there also were people of many other ancestries: Asian (4.6%), South East Asian (1.5%), Indian (1.5%), Australian Aboriginal (1.1%), Middle Eastern (1.0%), South Pacific (1.0%), African (0.4%), and South American (0.2%). Some members of the sample did not provide information on their ethnic background or were of mixed ancestry.

Procedure

Data were collected by 11 undergraduate honor's students (9 females and 9 of European ancestry) and a female research assistant of European ancestry. Target-present and target-absent lineups were produced for each of the 11 students. After dividing into six pairs, each pair was asked to select two locations for data collection such that there was a place where the students could remain out of sight until a participant was recruited and where viewing distance could be varied. Data were collected at two distances at each location. One distance fell in the range from 4 to 15 m while the other distance was in the range from 20 to 50 m. Locations were varied from on campus to city streets and a botanical garden.

When suitable people (individually or in small groups) appeared they were approached and asked if they were willing to donate a small amount of time to assist with some research for the university and if so whether they had an email address that they were willing to provide. Once the participant agreed, the target was signaled to come out of hiding. The participant's attention was directed to the target in such a way that it was clear who they were to look at and such that there were no people or other obstructions between the target and participant during the viewing (though often other people would be in the vicinity). The target remained facing the participant for approximately 10 s and then returned to hiding. Participants were not warned that they would have to identify the target or estimate their distance from the target. Targets were viewed in daylight conditions.

In the perceptual judgment condition, the participant was asked to estimate the distance to the target and respond to questions about themselves and the target while facing in the direction they had seen the target. Thus, participants could see the location where the targets were seen, but could not see the targets themselves. In the immediate judgment condition, the participants were asked to turn around and remain facing away from the direction of the target's location while responding to the questions. In both these judgment conditions the questions were answered on forms with a pen or pencil and clipboard provided by the researcher. In the delayed judgment condition the participants were told that an email would be sent by the next day and that they would complete the study via the internet. All participants answered the same questions in the same order.

Clearly the perceptual and immediate judgment conditions differ from the delayed condition both in timing and the availability of contextual cues to aid the distance judgments. This confound reflects the expected situation for many real-world witnesses who are, often, not likely to be at the scene of the crime if their judgments are solicited long after the event. Another difference in the delayed condition is that the photos were presented via computer. We matched the size and quality of the photos as presented on the computer as closely as we could to their size and quality as presented in the other two conditions.

Questions

Demographic questions determined the participant's age, sex, and ethnicity as well as whether the participant wore glasses and, if so, whether the glasses were on when viewing the target. Next the participant described the target in terms of age, sex, height, weight, and ethnicity. Participants also indicated if they knew the target (any that did were dropped from the sample). Participants then provided estimates of the distance between themselves and the target as well as estimates of the minimum and maximum distance they felt the target could have been from them. All distance estimates were made in units preferred by the participant. Distances originally were estimated in feet, yards, centimeters, and meters but all estimates were converted to meters prior to data analysis.

Identification Task

After completing the questions just described, participants were instructed to turn the page over (or to click a button to move to a new web page). On the back of the questionnaire (or the new web page), participants were asked to describe the target. Next they were shown a lineup. Written instructions informed the participant that "The person you saw may or may not be in the set of photographs you are about to examine." The page also contained a response form with the numbers from one to six arranged in two rows of three above the phrases "not there" and "not sure" that were separated by a noticeable space. The participant was then presented with a target-present or target-absent, six-person, simultaneous lineup. Fillers had been selected to be a reasonable but not exceptional match to the appearance of the target as judged by the co-author who constructed the lineups (CS). The photographs were presented in two rows of three. The participant was instructed orally and in writing to circle (click on) one response (number of the recognized photo, not there, or not sure). This completed the data collection. The participants were debriefed.

Design

People were randomly assigned to one of three judgment conditions: Perceptual ($n = 400$), immediate ($n = 430$), or delayed ($n = 491$) prior to attrition. Within each of the distance estimation conditions, approximately equal numbers of participants viewed each student target. Participants were approximately equally likely to be presented with a target-present versus target-absent lineup across distance estimation conditions, locations, and targets. Participants were tested at either a short or a long distance for each location. Long distances varied between 20 and 50 m ($M = 33.42$, $SD = 8.21$) and short distances varied between 5 and 15 m ($M = 9.97$, $SD = 3.22$). These distance groups were then used as the grouping variable in subsequent analyses. Effect size is reported as Cohen's w (cutoffs for small, medium, and large effects are .10, .30, and .50, respectively) for non-parametric tests and Cohen's f (cutoffs for small, medium, and large effects are .10, .25, and .40, respectively) for parametric tests.

Data Screening and Attrition

Attrition occurred in the delayed judgment condition due to participants failing to respond to the email or having given an incorrect email address ($n = 75$) leaving 416 participants in that condition. A small number of participants were eliminated because they knew the target ($n = 30$, six from the Perceptual condition, 11 from the Immediate, and 13 from the Delayed condition). The data were examined for outliers on the estimation of distance. Because of the large sample size, all data for participants whose distance estimates exceeded three standard deviations from the mean distance estimates for the actual distance they saw the target were eliminated from further consideration. This approach produced quite small changes to the cell sizes as data from only 20 participants were eliminated (13 from the Perceptual condition, 1 from the Immediate condition, and 6 from the Delayed condition). However, 15 participants provided no point estimate of the distance to the target and, of these, five provided no minimum and maximum distance estimate (none from the Perceptual condition leaving a total of 381, 2 from the Immediate condition leaving a total of 416, 13 from the Delayed condition leaving a total of 378). The final number of participants with complete data was 1,175. The distance estimates were significantly negatively skewed and transformation (using the square root, natural logarithm, and inverse of the estimates) did not improve the shape of the distribution enough to allow parametric tests of significance on raw distance estimates. Thus, comparisons are made using either relative or absolute error in distance estimates (see below for definitions).

Results

Relationship Between Actual and Estimated Distance

Distance estimates correlated significantly with actual distance in all three judgment conditions (Perceptual $r_s = .73$, $p < .001$; Immediate $r_s = .74$, $p < .001$; and Delayed $r_s = .67$, $p < .001$); however, this does not reflect a lack of error in the estimates. In absolute terms, errors in individual distance estimates varied from 0 to 700 m (for one of the outliers dropped from the analyses). The mean value of absolute error was as low as .05 m (immediate condition at 9–12 m) and as high as 12.57 m (delayed condition at 35–41 m). Variance in the distance estimates was larger when participants viewed the target at greater distances. Although this was expected it meant that the assumptions of ANOVA were not met for this variable. To overcome this problem, relative error was calculated by subtracting the estimated distance at which the target was viewed from the actual distance and dividing it by the actual distance. A 3 (judgment condition) $\times$ 7 (distance group) ANOVA on this variable revealed a small significant main effect of judgment condition, F (2, 1154) = 7.34, $p < .01$, $f = .11$, and small significant main effect of distance group, F (6, 1154) = 3.18, $p < .01$, $f = .13$. There was no interaction between judgment condition and distance condition, F (12, 1154) < 1, *ns*, $f = .09$.

Post hoc analysis (Tukey's HSD) showed that relative error was significantly greater in the delayed judgment condition than in either the immediate or perceptual conditions (see Table 1). Post hoc comparisons across distance

Table 1 Mean (and standard deviation) relative error[a] as a function of actual distance and judgment condition

	Distance group (m)							
	5–8	9–12	13–15	20–24	31–35	36–41	43–50	Overall
Perceptual	.10 (.45)	−.01 (.63)	.11 (1.18)	.20 (1.00)	.07 (.67)	−.14 (.58)	.21 (.76)	.07 (.78)
Immediate	.09 (.39)	−.01 (.56)	.09 (.83)	.03 (.58)	.01 (.62)	−.08 (.74)	−.04 (.65)	.03 (.61)
Delayed	.18 (.61)	−.17 (.43)	−.12 (.45)	−.25 (.44)	−.11 (.58)	−.32 (.43)	−.12 (.74)	−.11 (.54)
Overall	.12 (.48)	−.06 (.55)	.03 (.90)	.06 (.77)	−.02 (.62)	−.19 (.59)	.00 (.72)	−.01 (.66)
N	198	234	179	111	268	120	65	1,175

Notes: [a]Relative error = [(estimated distance − actual distance)/actual distance]

groups showed that the overestimates of distance were significantly greater at 5–8 m compared with all other distance groups. Underestimates were also significantly greater at 36–41 m compared will all other distance groups. Relative error in distance estimates did not take account of the fact that over- and underestimates cancel one another, and so it was also of interest to look at absolute relative error in distance estimates. A 3 (judgment condition) × 7 (distance group) ANOVA on absolute relative errors in distance estimates showed a main effect of distance group, F (6, 1154) = 3.34, $p < .01$, $f = .13$. There was no main effect of judgment condition, F (2, 1154) < 1, *ns*, $f = .05$, nor was there an interaction, F (12, 1154) < 1, *ns*, $f = .09$. Post hoc analysis on the main effect of distance group on absolute relative error showed that the error was smallest in the 5–8 and 9–12 m groups compared with all the other distance groups. Further, error was significantly smaller in the 31–35 m group than in the 43–50 m group. In relative absolute terms, individuals misestimated by as little as 0% and as much as 1,500% (again an outlier). Average relative absolute estimates were off by as little as 0.5% (in the Perceptual condition) and as much as 32.3% (underestimated in the Delayed condition).

Another important measure of the accuracy of distance estimates is the frequency of the actual distance falling within the limits provided by the witness. Even if a witness estimates a criminal was 10 m away when the true distance was 20 m, the witness may state that the point estimate is associated with a possible range from 5 to 30 m and thus the true value falls within the range provided. The proportion of “in range” estimates was examined across conditions. The size of the ranges provided varied considerably (from less than 1 to 200 m). The mean size of the range was significantly larger when targets were viewed at longer ($M = 20.93$, $SD = 21.69$) compared with shorter distances ($M = 6.12$, $SD = 6.21$), t (1184) = −15.72, $p < .001$. Overall, 47.8% of the actual distances fell within the range estimated by the participants. Although the percentage of within-range estimates differed significantly across distance, the pattern was not one of consistent decrease in accuracy with distance (64, 57, 37, 51, 42, 30, and 45%). A similar pattern was observed across the three judgment conditions (perceptual: 71, 55, 36, 57, 45, 31, and 47%; immediate: 66, 68, 39, 46, 46, 27, and 35%; and delayed: 55, 46, 36, 50, 37, 32, and 50%), with percentages slightly lower overall in the delayed judgment condition.

Description Accuracy and Distance

Data were collected about the targets’ height, weight, and age. Insisting on a precise match is clearly too stringent a criterion for accuracy; so the height, weight, and age estimates were treated as accurate if they fell within a specified range above and below the actual values associated with the targets. Because the size of these ranges is arbitrary, analyses were performed with accuracy defined as within (plus or minus) 2, 5, or 10 cm, kg, and years. Hierarchical loglinear models were used to test the possibility that increased distance reduces the accuracy of description. For estimates of age, there was no significant relationship between distance condition (long or short) and the percentage of estimates within 2, 5, or 10 years of the actual ages of the targets, χ^2 (2, N = 1,190) < 1, *ns*, (see Table 2). Estimates of weight were related to the distance at which the targets were viewed, with significantly more estimates falling within the 2- and 5-kg ranges when targets were viewed from shorter distances, χ^2 (2, N = 1,190) = 4.12, $p < .05$, $w = .06$, and χ^2 (2, N = 1,190) = 3.99, $p < .05$, $w = .06$, for 2 and 5 kg, respectively. There was no significant relationship between distance and weight estimates for the 10-kg range, χ^2 (2, N = 1,190) = 1.90, $p > .05$, $w = .04$. For height estimates there was a significant relationship with distance for all three ranges, χ^2 (2, N = 1,190) = 4.01, $p < .05$, $w = .06$, χ^2 (2, N = 1,190) = 4.61, $p < .05$, $w = .06$, and, χ^2 (2, N = 1,190) = 4.65, $p < .05$, $w = .06$, for 2, 5, and 10 cm, respectively. Patterns were very similar across the perceptual and delayed judgment conditions, with reduced accuracy overall in the delayed condition; for example,

Table 2 Percent of correct estimates (and frequency) of height, weight, and age (within 2, 5, and 10 cm, kg, and years) as a function of distance

Description	Height		
	Within 2 cm*	Within 5 cm*	Within 10 cm*
Distance condition			
Long	12.2 (75)	33.0 (202)	59.8 (366)
Short	19.0 (107)	45.3 (255)	76.6 (431)
Overall	15.5 (182)	38.9 (457)	67.8 (797)
	Weight		
	Within 2 kg*	Within 5 kg*	Within 10 kg
Long	17.1 (105)	38.6 (236)	61.4 (376)
Short	25.4 (143)	51.5 (290)	76.0 (428)
Overall	21.1 (248)	44.8 (526)	68.4 (804)
	Age		
	Within 2 years	Within 5 years	Within 10 years
Long	28.4 (174)	41.1 (252)	43.0 (263)
Short	32.3 (182)	47.5 (268)	49.6 (279)
Overall	30.3 (356)	44.3 (520)	46.1 (542)

* Values within column differ at .05

there was a significant relationship between accuracy of age estimates within 2 years and judgment condition, χ^2 (2, $N = 1{,}190$) = 108.85, $p < .001$, $w = .29$, with more estimates falling outside the range in the delayed judgment condition (89%) than the perceptual (62%) and immediate (60%) judgment conditions (more detailed data may be obtained from the third author on request).

Identification Responses

We were interested in three questions with regard to identification responses. First, does viewing distance affect the number of 'not sure' responses given by witnesses? We hypothesized that witnesses should be more likely to respond 'not sure' after viewing targets from longer distances compared with shorter distances, if they used viewing distance as a cue to determine whether or not they should attempt an identification. Looking first at target-absent lineups, a 2 (distance: under or over 15 m) × 3 (judgment condition: Perceptual, Immediate, or Delayed) × 3 (identification response: choice, not there, or not sure) hierarchical loglinear analysis revealed no significant interaction between identification response and distance group, χ^2 (6, $N = 560$) = 7.13, $p > .05$, $w = .11$ (see Table 3). Apparently, the pattern of identification responses for target-absent lineups was no different when targets were viewed at long as compared with short distances. The Identification Response × Judgment Condition interaction was not significant, χ^2 (6, $N = 560$) = 8.11, $p > .05$, $w = .11$. An anonymous reviewer pointed out that participants were not randomly assigned to the targets and lineups viewed within each distance group, presenting the possibility that any significant effects of distance may be due to differences in the distinctiveness of targets or the functional size of lineups rather than being due to distance per se. To explore this problem we examined the relationship between distance group, judgment condition, and identification response, with each of the targets excluded, to determine whether the pattern of results changed with the exclusion of responses to each target. The pattern of results does not change with the exclusion of each target. Short distances produced greater accuracy in the target-present conditions for all but one (of 22) tests and no significant difference in the target-absent comparisons for all 22 tests.

For target-present lineups, a 2 (distance group) × 3 (judgment condition) × 4 (identification response: target selection, filler selection, not there, or not sure) hierarchical loglinear analysis revealed a significant Identification Response × Distance Group interaction, χ^2 (3, $N = 554$) = 36.01, $p > .001$, $w = .22$, with the percentage of 'not sure' responses doubling in the long distance condition (from 7.02% to 15.22%). Apparently, responses to target-present lineups were affected by viewing distance. We were also interested in determining whether the rate of choosing decreased as distance increased. This should be reflected in a reduction in the number of suspect and foil choices at longer distances. Suspect identifications were substantially lower for long compared with short distances and the percentage of foil identifications was fairly similar across the distance groups (see Table 3). There was also a significant Identification Response × Viewing Condition interaction, χ^2 (6, $N = 554$) = 19.97, $p < .01$, $w = .17$, with a greater percentage of 'not sure' responses in the delayed judgment condition (44%) than in the perceptual, or immediate judgment conditions (29 and 26%, respectively). This was also the case for incorrect rejections (40, 30, and 30% for delayed, perceptual, and immediate conditions, respectively). The percentage of correct identifications also varied across judgment condition (27, 34, and 38% for delayed, perceptual, and immediate conditions, respectively). Consistent with the higher rates of incorrect rejection and 'not sure' responses, there also were far fewer incorrect identifications in the delayed condition (18%) than in the perceptual (39%) and immediate (42%) judgment conditions. No other significant effects were observed for either target-absent or target-present lineups.

A different way of looking at the effects of distance and judgment condition on identification performance is to consider their effect on the diagnosticity ratio. Wells and

Table 3 Identification response (%) as a function of target presence and distance and judgment conditions

Target presence[a]	Distance condition					
	Short			Long		
	Target-present					
View condition	Percep	Immed	Delayed	Percept	Immed	Delayed
ID suspect	61.3	63.3	58.6	34.9	37.2	38.2
ID foil	20.8	21.1	14.3	39.5	39.4	17.9
Incorrect rejection	11.3	9.2	18.6	12.8	13.8	20.2
Not sure	6.6	6.4	8.6	12.7	9.6	23.6
N	106	109	70	86	94	89
	Target-Absent					
View condition	Percep	Immed	Delayed	Percept	Immed	Delayed
Incorrect ID	57.3	45.5	43.0	58.3	55.7	45.9
Correct rejection	35.4	42.7	47.1	26.2	30.8	39.6
Not sure	7.3	11.8	9.9	15.5	13.4	14.4
N	96	110	121	84	97	111
Diagnosticity[b] ratio (choosers)	6.42	8.35	8.18	3.59	4.01	4.99
Diagnosticity ratio (non-choosers)	3.13	4.64	2.53	2.05	2.23	1.96

[a] Lineup type (target-absent or target-present) was not recorded for 17 participants

[b] Diagnosticity of choosers (Wells and Lindsay 1980) is the ratio of the correct identification rate to the false identification rate. False identification rates were estimated by dividing the false positive choice rate from target-absent lineups by the nominal size of the lineup (6). Diagnosticity of non-choosers is the ratio of correct rejection rates to incorrect rejection rates

Lindsay (1980) defined the diagnosticity ratio as the ratio of (the rates of) correct to false suspect identifications. The higher the ratio, the more likely an identified suspect is to be the criminal. Despite the fact that the rate of correct choosing declined by approximately 25% as distance increased, the reduction in the diagnosticity ratio from the short to the long distance group was 35% (see Table 3).

Witness Characteristics

We also examined the effect of participant age on identification response. Participants were divided into three groups according to their age (a total of 1,157 provided age information): 18–23 year olds ($n = 653$), 24–49 year olds ($n = 391$), and 50 and above ($n = 113$). There was no significant relationship between age and identification response for either target-absent, χ^2 (4, $N = 607$) = 6.71, *ns,* $w = .11$, or target-present, χ^2 (6, $N = 550$) = 6.39, *ns,* $w = .11$, lineups. However, the effect sizes for both analyses are indicative of small effects. We also examined the relationship between identification response and the ethnicity of the targets and witnesses. Based on the reported ethnicity of the participant and the ethnicity of the target, each identification decision was classified as either own- or other-ethnicity and a 2 (ethnicity) × 2 (identification accuracy) × 2 (choosing status) hierarchical loglinear analysis was conducted. It revealed no evidence for a significant impact of ethnicity on either identification accuracy, χ^2 (1, $N = 1010$) = 1.88, *ns*, $w = .06$, or choosing status, χ^2 (1, $N = 1010$) = 2.27, *ns*, $w = .06$, and no evidence of a significant three-way interaction, χ^2 (1, $N = 1010$) = 0.11, *ns*. Furthermore, the negligible effect sizes (i.e., below 0.1, the criterion for a small effect) for both two-way interactions, along with the large sample sizes, indicate that these null findings were not the result of a lack of power.

Identification Response and Estimated Distance

As suggested earlier, accuracy may vary with either actual or perceived distance. To test the impact of perceived distance, the distance estimation data were used to divide the sample in a fashion similar to the way actual distance had been used. The seven groups generated estimated their distance from the targets as approximately 5, 10, 15, 20, 30, 40, and 65 m. Distance estimates shorter than 30 m tended to be underestimates of the true mean distance, while estimated distances over 30 m tended to be overestimates of actual distance (see Table 4). The accuracy of correct selections varied significantly across the perceived distance conditions, χ^2 (6, $N = 727$) = 23.23, $p < .001$,

Table 4 Mean estimated distance (m), mean actual distance (m), estimated distance as a percent of actual distance, correct identification rate, and correct rejection rate of seven groups formed based on estimated distance

N	*M* estimated distance	*M* actual distance	Est. as % of actual	% of correct ID	% of correct rejection
333	5.50 (1.68)	10.71 (7.58)	51.4	66.0	45.1
232	10.30 (0.87)	14.64 (9.05)	70.4	55.2	37.9
121	14.74 (0.67)	21.55 (11.98)	68.4	40.4	32.8
142	19.64 (1.34)	26.48 (10.95)	74.1	34.8	44.6
178	28.53 (2.94)	31.15 (9.36)	91.5	38.7	35.7
36	39.81 (0.82)	36.12 (7.69)	110.2	31.3	20.0
133	65.94 (24.71)	34.63 (10.33)	190.4	43.8	25.9

Note: Est. = Estimated distance

$w = .26$. As perceived distance increased, correct identifications tended to decline, though, as shown in Table 4, the result was not a pattern of steady decreases. Similarly, correct rejections were significantly related to perceived distance but the pattern was not one of consistent decreases as distance increased, χ^2 (6, $N = 685$) = 14.46, $p < .05$, $w = .21$.

Although participants, particularly in the delayed condition, demonstrated considerable error in their judgments of distance to the target, these judgments of perceived distance were still related to accuracy. The key question for the courts then is: do perceived and actual distance differ in their usefulness as indicators of accuracy? To answer this question we conducted, separately for choosers and non-choosers, logistic regressions with accuracy as the dependent measure. Perceived distance was entered on the first step and actual distance the second. For non-choosers, neither perceived distance alone, χ^2 (1) = 1.62, *ns*, nor perceived and actual distance together, χ^2 (2) = 3.49, *ns*, significantly predicted the accuracy of the rejection decision. In contrast, for choosers, perceived distance did significantly predict accuracy, χ^2 (1) = 13.20, $p < .01$. Further, when actual distance was added, the model was a significantly better fit to the data, χ^2 (1) = 19.53, $p < .01$, and perceived distance was no longer a significant predictor of accuracy, $B = -0.00$, odds ratio = 1.00, Wald (1) = 18.59, *ns*. Examination of Nagelkerke R^2 values (perceived distance only: $R^2 = .03$; perceived and actual distance: $R^2 = .06$) indicates that the strength of the relationship between identification accuracy and the distance variable (or variables) was twice as strong when actual distance was included in the model. Thus, perceived distance was a significant predictor of accuracy, but actual distance was a substantially stronger predictor.

Discussion

One issue addressed in this study was the accuracy of witnesses' judgments of the distance of a person they had observed for a brief period of time. Previous research demonstrates that people have difficulty judging the distance between themselves and inanimate objects and between two inanimate objects (Wiest and Bell 1985; Witmer and Kline 1998; Wright 1995). Our data extend this conclusion to judgments of the distance between oneself and another person. Errors of judgment were often substantial, with the delayed judgment condition—the condition that arguably most closely approximates the situation for real-life witnesses—producing underestimates. The range of distance estimates was typically quite large, especially when considered in relation to the actual distance. Moreover, when witnesses gave a range of distances rather than a point estimate, the range frequently did not contain the actual distance. Taken together, these data clearly demonstrate that it would be very difficult for the courts to establish what the actual distance between a witness and the criminal had been solely on the basis of witness estimates.

Second, we explored the impact of distance on the reliability of witness descriptions. Given the sizable errors associated with distance estimates it is perhaps surprising that the reliability of witnesses' descriptions (at least for height, weight, and age) was virtually unaffected across distances up to 50 m. Nevertheless, it should be noted that the grain size of witness reports that was necessary to capture the 'true' value for each descriptor was quite broad for the height and weight dimensions, implying that estimates on these descriptors should only be considered to be relatively rough.

The third, and the major, issue examined was the variation of identification accuracy with distance. Accuracy of witness identification decisions was significantly influenced by the distance between the witness and the target at the time of exposure. The decline in accuracy occurred both for target-present and target-absent lineups, though the effect size was larger when the target was present. This pattern of results was both intuitively likely and consistent with previous research (De Jong et al. 2005; Wagenaar and van der Schrier 1996). Based on their findings, Wagenaar and his colleagues suggested a 15-m rule as a useful "rule of thumb" for the courts. What can such a rule mean? If the courts accepted this reasoning and instituted an exclusionary rule, identification evidence produced by witnesses who viewed a crime from more than 15 m would not be permitted as evidence. This approach contains the questionable assumption that identification accuracy changes dramatically at about 15 m such that performance is acceptably high for shorter distances but unacceptably low for greater distances. Our data demonstrate that this is not a reasonable assumption. The accuracy of choosers declined with distance but the decline occurred both under and over 15 m. Furthermore, there was no noticeable increase in the rate of decline after as opposed to before the 15-m distance. As a result, the distance between the witness and criminal should "go to the weight" of the evidence, not its admissibility. A less stringent approach would suggest that the evidence could be presented but that the weight accorded to identification evidence obtained following an exposure from more than 15 m should be minimal. Once again, this does not align well with our data. Even at 43 m, identification evidence has some diagnostic value, and therefore probative value as well. Less weight should be accorded to identification evidence as distance increases, but further research would be required to determine the distance and conditions that reduce diagnosticity to one and thus probative value to zero. The answer to one of the issues we set out to address is clear: the 15-m rule is not particularly useful for the courts.

Our speculations regarding "not sure" responses were not supported. Although participants shown target-absent lineups did increase their tendency to say "not sure" as distance increased, the change in the "not sure" rate was non-significant despite the large sample size. This is a disturbing finding and may support other recent contentions that witnesses are likely to guess when presented with simultaneous lineups (e.g., Lindsay et al. in press; Penrod 2006).

Two other features of our data warrant emphasis. First, we had speculated that people may infer distance from the strength of their memory trace for the person. If true, estimated distance, even if in error, may predict identification accuracy. Accuracy of both choosers and non-choosers differed significantly among the perceived distance groups. However, examination of the accuracy rates and logistic regressions indicates different patterns for choosers and non-choosers. Consistent with results for other potential indicators of identification accuracy (e.g., confidence: Brewer and Wells 2006; and response latency: Weber et al. 2004), neither perceived nor actual distance was systematically related to the accuracy of non-choosers. Thus, perhaps counter-intuitively, a lineup rejection made by a witness who viewed the target from 10 m is no more likely to be correct (or incorrect) than a rejection by a witness 30 m from the target. In contrast, the accuracy of choosers declined with increasing distance, both perceived and actual. Importantly, the predictive strength of actual and estimated distance differed substantially, with actual distance being the superior predictor. Consequently, objective evidence of actual distance is important for the courts. However, when such evidence is not available, perceived distance must serve as a proxy and will not be entirely uninformative.

Second, we investigated two factors that have been consistently demonstrated to influence identification performance in laboratory studies, namely, witness age (e.g., Searcy et al. 1999, 2001) and the match between the ethnicity of the witness and of the target (i.e., the cross-race or other-race effect, e.g., Meissner and Brigham 2001). Interestingly, we found no evidence for significant effects of either of these factors on identification performance. One explanation for this null finding is that our sampling led to uneven marginal frequencies for categories of age and of target-witness ethnicity. As analyses of categorical data are less powerful when marginal frequencies diverge, this imbalance may have reduced our chances to detect these effects. However, examination of the effect-size measures, which were of negligible size for target-witness ethnicity and barely of small size for age, indicates that these null findings were not the sole result of a lack of power. Another limitation of the cross-race comparison is the limited number of targets such that they may not be representative of their ethnic groups (particularly the non-Europeans). However, our number of European targets was much greater than the typical number found in most staged event studies and our sample size also is extremely large. If the cross-race effect is robust, one would have expected it to occur under the conditions of our study.

Further, these slight effect sizes suggest a more satisfying explanation of our failure to detect significant effects of age and target-witness ethnicity. In a typical laboratory experiment the stimuli and procedure are carefully controlled to reduce extraneous variance in identification performance. In contrast, in our field study we were unable to control many such factors and, indeed, deliberately introduced variance in a number of others (i.e., location, viewing distance, retention interval, the target, and the lineup). In other words, our study included more methodological sources of variance than laboratory studies investigating the impact of age and ethnicity on identification performance and, therefore, likely observed greater variance in identification performance. As a result, the error variance, against which the variance accounted for by age and ethnicity was compared, was potentially substantially larger in our study. Thus, a plausible interpretation of our results is that, although there are readily apparent effects in the laboratory, the amount of variance accounted for by these effects is not substantial when compared with the natural variance in identification performance likely to be observed in a more ecologically valid sample of identification decisions. This explanation shares parallels with the observation that the artificially low variance in performance typical of identification experiments reduces the confidence-accuracy relationship (Lindsay et al. 2000; Lindsay et al. 1998) and has important implications for the study of eyewitness identification. Specifically, although it is important to reduce extraneous variance when attempting to understand the impact of a factor or combination of factors on identification performance, such an approach is not a valid tool for assessing the size of the effect of these factors on real-world performance. In other words, to assess the extent to which a factor is likely to affect real-world identification performance, it is important to compare explained variance with a valid estimate of the variance in performance of real-world identifications.

In conclusion, let us return for a last time to the 15-m rule. Wagenaar et al. used a ratio of 5:1 of correct to false identifications (or a diagnosticity ratio of five) to determine the distance that defined a cut-off for acceptable identification evidence. However, the 5:1 ratio is arbitrary. Our data confirm that identification for all distances under 15 m produced diagnosticity ratios above 5 and all distances over 15 m produced diagnosticity ratios below 5. Still, the decline in accuracy is not dramatic at 15 m. Rather correct identifications decrease with increasing distance in general. Of greater concern is the failure of correct rejections to increase with distance. Instead, many witnesses who were too far from the target to permit an accurate identification decision chose to guess.

References

Brewer, N., & Wells, G. L. (2006). The confidence-accuracy relationship in eyewitness identification: Effects of lineup instructions, foil similarity and target-absent base rates. *Journal of Experimental Psychology: Applied, 12*, 11–30.

De Jong, M., Wagenaar, W. A., Wolters, G., & Verstijnen, I. M. (2005). Familiar face recognition as a function of distance and illumination: A practical tool for use in the courtroom. *Psychology, Crime & Law, 11*, 87–97.

Greene, E., & Fraser, S. C. (2002). Observation distance and recognition of photographs of celebrities' faces. *Perceptual & Motor Skills, 95*, 637–651.

Harway, N. I. (1963). Judgment of distance in children and adults. *Journal of Experimental Psychology, 65*, 385–390.

Illinois v Levin 1991, 152 Ill. App. 824.

Kerst, S. M., & Howard, J. H. (1978). Memory psychophysics for visual area and length. *Memory & Cognition, 6*, 327–335.

Kunnapas, T. (1960). Scales of subjective distance. *Scandinavian Journal of Psychology, 1*, 187–192.

Lindsay, D. S., Nilsen, E., & Read, J. D. (2000). Witnessing-condition heterogeneity and witnesses' versus investigators' confidence in the accuracy of witnesses' identification decisions. *Law & Human Behavior, 24*, 685–697.

Lindsay, D. S., Read, J. D., & Sharma, K. (1998). Accuracy and confidence in person identification: The relationship is strong when witnessing conditions vary widely. *Psychological Science, 9*, 215–218.

Lindsay, R. C. L., Mansour, J. K., Beaudry, J. L., Leach, A. M., & Bertrand, M. I. (in press). Simultaneous and sequential lineups: The problem of eyewitness identification error. *Legal & Criminological Psychology*.

Loftus, G. R., & Harley, E. M. (2005). Why is it easier to identify someone close than far away? *Psychonomic Bulletin and Review, 12*, 43–65.

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, & Law, 7*, 3–35.

Penrod, S. (2006). *Eyewitness guessing and choosing.* Paper presented at the meeting of the American Psychology-Law Society, St. Petersburg, FL.

Searcy, J. H., Bartlett, J. C., & Memon, A. (1999). Age differences in accuracy and choosing in eyewitness identification and face recognition. *Memory & Cognition, 27*, 538–552.

Searcy, J. H., Bartlett, J. C., Memon, A., & Swanson, K. (2001). Aging and lineup performance at long retention intervals: Effects of metamemory and context reinstatement. *Journal of Applied Psychology, 86*, 207–214.

Stevens, S. S. (1957). On the psychophysical law. *Psychological Review, 64*, 153–181.

Thorndyke, P. W. (1981). Distance estimation from cognitive maps. *Cognitive Psychology, 13*, 526–550.

Wagenaar, W. A., & van der Schrier, J. H. (1996). Face recognition as a function of distance and illumination: A practical tool for use in the courtroom. *Psychology, Crime & Law, 2*, 321–332.

Weber, N., Brewer, N., Wells, G. L., Semmler, C., & Keast, A. (2004). Eyewitness identification and response latency: The unruly 10–12 second rule. *Journal of Experimental Psychology: Applied, 10*, 139–147.

Wells, G. L., & Lindsay, R. C. L. (1980). On estimating the diagnosticity of eyewitness nonidentifications. *Psychological Bulletin*, *88*, 776–784.

Wiest, W. M., & Bell, B. (1985). Steven's exponent for psychophysical scaling of perceived, remembered, and inferred distance. *Psychological Bulletin, 98*, 457–470.

Witmer, B. G., & Kline, P. B. (1998). Judging perceived and transversed distance in virtual environments. *Presence, 7*, 144–167.

Wright, D. B. (1995). Misinformation methodologies: Explaining the effect of errant information. In G. M. Davies, S. Lloyd-Bostock, M. McMurren, & J. C. Wilson (Eds.), *Psychology, law and criminal justice: International developments in research and practice* (pp. 39–45). Berlin: De Gruyter.