APPLIED COGNITIVE PSYCHOLOGY
Appl. Cognit. Psychol. **17**: 1067–1080 (2003)
Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/acp.986

Reconstructing the Serial Order of Events: A Case Study of September 11, 2001

ERIK M. ALTMANN*

Michigan State University

SUMMARY

Participants reconstructed the serial order of events associated with the terrorist attacks of September 11, 2001. Accuracy was stable as the retention interval doubled from 9 weeks to 18 weeks, but most participants (63%) made errors sequencing even the 'unforgettable' events of that day, with most errors (72%) falling within one position of being correct. This *positional gradient* (near misses more common than far misses) mirrored that found in laboratory studies of order memory, despite basic differences in how ordinal information was presented to observers, suggesting that the positional gradient may generalize across many different circumstances of serial-order reconstruction. Monte Carlo simulations illustrate this possibility by showing that simple constraints on pairs of events explain substantial variance in the data. Copyright © 2003 John Wiley & Sons, Ltd.

This article examines how well people were able to reconstruct the serial order of events associated with the terrorist attacks in the United States on September 11, 2001, and asks what the resulting error patterns tell us about memory representations that support order reconstruction in real-world situations. To motivate the interest in serial order reconstruction, as distinct from more general tests of recall, consider a hypothetical situation in which a witness observes an altercation between two individuals and is later asked to testify in an inquiry in which one party is claiming self-defense. Conceivably, the focus of the inquiry might not be on *whether* a fight occurred, but on whether blame can be assigned by identifying the person who struck *first*. As a second example, consider the inquiry into the crash of EgyptAir flight 990 off the coast of Massachusetts in 1999 (NTSB, 2002). Two very different hypotheses were proposed for the cause of the crash—mechanical vs. psychological—and deciding between them depended on correctly reconstructing the serial order of two events: The onset of the fatal dive into the ocean, and the onset of the pilot's prayers commending his life to God. These examples illustrate that the relative order of two events can be a critical factor in causal reasoning in a wide range of circumstances. To the extent that human memory is a source of information about serial order of events, it is important to understand how such memory functions.

*Correspondence to: Erik M. Altmann, Department of Psychology, Michigan State University, East Lansing, Michigan 48824, USA. E-mail: ema@msu.edu

Existing data on serial-order reconstruction come primarily from laboratory studies, where the 'events' are typically random lexical items chosen so as to have no obvious relation to one another, and where the only information about the ordinal relationship between the events is conveyed by having participants experience them in a particular sequence. In a representative example of such studies, Nairne (1992) presented participants with lists of random words, at a rate of one word every 3 s, with the nominal task being to rate the 'pleasantness' of each word. Participants heard lists of five words each, then after a retention interval (ranging from 30 sec to 24 hours) were given a surprise memory test in which they were shown the same words in random order and asked to reconstruct the serial order in which the words were originally presented. Order reconstruction was reasonably accurate, in that the modal placement for each item was its original position, even after a 24-hour retention period. However, perhaps the more interesting result was that when an item was sequenced incorrectly, it was more likely to be placed nearer to its original position than farther away. (This pattern is shown in the bottom panels of Figure 2, and will be discussed again later in the article.) This *positional gradient* has been interpreted to mean that event sequences have an explicit temporal representation in memory, which may distort as a function of noise in the system but nonetheless preserves basic ordinal or temporal locality (G. D. A. Brown, Preece, & Hulme, 2000; Estes, 1997; Page & Norris, 1998). The positional gradient implies a second pattern, the *serial position* curve, in which items at the ends of a list are sequenced more accurately than items in the middle. The standard interpretation of the serial position curve (in order-memory experiments) is that end items can 'drift' out of sequence in only one direction (towards the middle of the list), whereas middle items can drift in two directions (toward either end), making them more prone to error.

In standard order-memory experiments, the 'events' are stylized in various ways compared to events one might experience in the real world, but there is nonetheless some reason to expect serial-order reconstruction of real-world events to exhibit similar patterns. For example, a positional gradient arises when people report the day of the week on which they remember an event taking place, with most errors falling on a day neighboring the correct one (Huttenlocher, Hedges, & Prohaska, 1992). Also, a temporal gradient arises when people are asked to assign dates to events (Wagenaar, 1986), though this gradient is often asymmetrical (Huttenlocher, Hedges, & Bradburn, 1990; Loftus & Marburger, 1983; Shum, 1998). Nonetheless, few studies of autobiographical memory report results specifically of serial-order reconstruction tasks in which events are presented in random order and the task is to sequence them. Thus, one goal of the current study is to address the empirical question of whether the positional gradient arises in sequencing sets of real-world events. Two related empirical aims are to estimate, for this one case study, how rarely or commonly people err in sequencing a set of events consensually viewed as 'unforgettable,' and to assess the role of retention interval on memory for the information that supports serial order reconstruction.

If serial-order reconstruction of events of September 11 does exhibit a positional gradient, this would raise theoretical questions concerning the underlying memory representation, given the numerous differences between these events and those encountered in the usual order-memory experiment. Differences include the salience of the events (momentous vs. mundane), the rate at which events are presented (tens of minutes vs. a few seconds), and the range of inter-event intervals (tens of minutes vs. tenths of a second). With respect to theoretical models of the positional gradient (G. D. A. Brown et al., 2000; Estes, 1997; Page & Norris, 1998), perhaps the most critical difference is that in the experimental setting the observer directly experiences the time course of events as they unfold. According to these models, the system maps each event, as it is experienced, to a distinct internal context in a sequence of contexts generated by an internal process in some replicable way. At test time, the system then 'replays' this internal ordinal sequence, with each context in the sequence indexing the associated event. Adding noise to this process, such that neighboring contexts are occasionally confused, directly produces the positional gradient. Such *ordinal-context* explanations of the positional gradient make sense given the nature of the stimuli in order-memory experiments. Because these are typically chosen so as to appear random and unrelated, their time course is the *only* ordinal information provided to the system, which must therefore preserve this information with its own, replicable ordinal representation to have any success at sequencing the events later.

In the real-world setting, in contrast, ordinal information is often coded descriptively, such that the observer must extract it through semantic or inferential processing of some kind. For example, Figure 1 shows two photographs from the New York *Times* capturing the moments surrounding the collision of the second hijacked airplane into the World Trade Center on September 11. (The larger of the two images appeared again in the December 31, 2001 edition of the *Times*, occupying the full front page of 'The Year in Pictures' section.) In both photographs, and in several others like them, the first tower is clearly still standing, though also clearly on fire. Such images support the *inference* that both airplanes collided before either tower collapsed, but do not constitute an experience of one event occurring before the other. Similarly, consider a declarative statement of the sort one might hear on the news: 'The crash at the Pentagon occurred 40 minutes after the second airplane struck the World Trade Center.' This statement correctly (semantically) indicates that the Pentagon was struck after both towers of the World Trade Center were struck, but again provides no experience of one event occurring before the other. (Indeed, if this is the observer's first 'experience' of the events, then he or she 'experiences' the Pentagon crash first, by virtue of it occurring first in the statement.) Thus, models of the positional gradient in order-memory experiments would not be the likeliest explanation of a positional gradient in the case of September 11. With widespread publication of images in which both towers are standing as the second collision occurs (Figure 1), and of rehearsal of the day's events in declarative statements like that above, descriptive representations may dominate—particularly at longer retention intervals—whatever ordinal-context representations might have been acquired through experiencing the events directly. Thus, a second goal of this study is to ask in theoretical terms how a positional gradient might arise when ordinal information cannot be 'read out' directly from an ordinal memory representation, but instead has to be inferred from descriptive representations. The basic notion, elaborated in the Discussion using Monte Carlo simulations, is that relatively simple logical and other constraints relating subsets of events, of the kind typically present when events are part of a narrative structure rather than entirely independent, can be sufficient to produce the positional gradient.

In the study described below, participants sequenced two sets of events that took place on September 11, 2001. One set involved the major events of the day—the hijacked airplanes crashing into various targets, and particular targets collapsing. A second set involved events that might be less familiar—the stops of the President that day, as he flew around the country. The empirical question was whether serial-order reconstruction within each set would exhibit a positional gradient, with an ancillary aim to assess the frequency with which errors occur in sequencing 'unforgettable' events.

"All the News That's Fit to Print"

The New York Times

Late Edition

New York: **Today,** sunny, a few afternoon clouds. High 77. **Tonight,** slightly more humid. Low 65. **Tomorrow,** sun then clouds. High 81. **Yesterday,** high 81, low 63. Weather map, Page C18

VOL. CL . . No. 51,874 — NEW YORK, WEDNESDAY, SEPTEMBER 12, 2001 — 75 CENTS

U.S. ATTACKED

HIJACKED JETS DESTROY TWIN TOWERS AND HIT PENTAGON IN DAY OF TERROR

A CREEPING HORROR

Buildings Burn and Fall as Onlookers Search for Elusive Safety

By N. R. KLEINFIELD

It kept getting worse.

The horror arrived in episodic bursts of chilling disbelief, signified first by trembling floors, sharp eruptions, cracked windows. There was the actual unfathomable realization of a gaping, flaming hole in first one of the tall towers, and then the same thing all over again in its twin. There was the merciless sight of bodies helplessly tumbling out, some of them in flames.

Finally, the mighty towers themselves were reduced to nothing. Dense plumes of smoke raced through the downtown avenues, coursing between the buildings, shaped like tornadoes on their sides.

Every sound was cause for alarm. A plane appeared overhead. Was another one coming? No, it was a fighter jet. But was it friend or enemy? People scrambled for their lives, but they didn't know where to go. Should they go north, south, east, west? Stay outside, go indoors? People hid beneath cars and each other. Some contemplated jumping into the river.

For those trying to flee the very epicenter of the collapsing World Trade Center towers, the most horrid thought of all finally dawned on them: nowhere was safe.

For several panic-stricken hours yesterday morning, people in Lower Manhattan witnessed the inexpressible, the incomprehensible, the unthinkable. "I don't know what the gates of hell look like, but it's got to be like this," said John Maloney, a security director for an Internet firm in the trade center. "I'm a combat veteran, Vietnam, and I never saw anything like this."

The first warnings were small ones. Blocks away, Jim Farmer, a film composer, was having breakfast at a small restaurant on West Broadway. He heard the sound of a jet. An odd sound — too loud, it seemed, to be

Continued on Page A7

A Somber Bush Says Terrorism Cannot Prevail

By ELISABETH BUMILLER with DAVID E. SANGER

WASHINGTON, Sept. 11 — President Bush vowed tonight to retaliate against those responsible for today's attacks on New York and Washington, declaring that he would "make no distinction between the terrorists who committed these acts and those who harbor them."

"These acts of mass murder were intended to frighten our nation into chaos and retreat, but they have failed," the president said in his first speech to the nation from the Oval Office. "Our country is strong. Terrorist acts can shake the foundation of our biggest buildings, but they cannot touch the foundation of America."

His speech came after a day of trauma that seems destined to define his presidency. Seeking to at once calm the nation and declare his determination to exact retribution, he told a country numbed by repeated scenes of carnage that "these acts shattered steel, but they cannot dent the steel of American resolve."

Mr. Bush spoke only hours after returning from a zigzag course across the country, as his Secret Service and military security teams moved him from Florida, where he woke up this morning expecting to press for his education bill, to command posts in Louisiana and Nebraska before it was determined the attacks had probably ended and he could safely return to the capital.

It was a sign of the catastrophic

Continued on Page A4

AMERICAN TARGETS A ball of fire exploded outward after the second of two jetliners slammed into the World Trade Center; less than two hours later, both of the 110-story towers were gone. Hijackers crashed a third airliner into the Pentagon, setting off a huge explosion and fire.

President Vows to Exact Punishment for 'Evil'

By SERGE SCHMEMANN

Hijackers rammed jetliners into each of New York's World Trade Center towers yesterday, toppling both in a hellish storm of ash, glass, smoke and leaping victims, while a third jetliner crashed into the Pentagon in Virginia. There was no official count, but President Bush said thousands had perished, and in the immediate aftermath the calamity was already being ranked the worst and most audacious terror attack in American history.

The attacks seemed carefully coordinated. The hijacked planes were all en route to California, and therefore gorged with fuel, and their departures were spaced within an hour and 40 minutes. The first, American Airlines Flight 11, a Boeing 767 out of Boston for Los Angeles, crashed into the north tower at 8:48 a.m. Eighteen minutes later, United Airlines Flight 175, also headed from Boston to Los Angeles, plowed into the south tower.

Then an American Airlines Boeing 757, Flight 77, left Washington's Dulles International Airport bound for Los Angeles, but instead hit the western part of the Pentagon, the military headquarters where 24,000 people work, at 9:40 a.m. Finally, United Airlines Flight 93, a Boeing 757 flying from Newark to San Francisco, crashed near Pittsburgh, raising the possibility that its hijackers had failed in whatever their mission was.

SECOND PLANE United Airlines Flight 175 nearing the trade center's south tower.

There were indications that the hijackers on at least two of the planes were armed with knives. Attorney General John Ashcroft told reporters in the evening that the suspects on Flight 11 were armed that way. And Barbara Olson, a television commentator who was traveling on American Flight 77, managed to reach her husband, Solicitor General Theodore Olson, by cell phone and to tell him that the hijackers were armed with knives and a box cutter.

In all, 266 people perished in the four planes and several score more were known dead elsewhere. Numerous firefighters, police officers and other rescue workers who responded to the initial disaster in Lower Manhattan were killed or injured when the buildings collapsed. Hundreds were treated for cuts, broken bones, burns and smoke inhalation.

But the real carnage was concealed for now by the twisted, smoking, ash-choked carcasses of the twin towers, in which thousands of people used to work on a weekday. The collapse of the towers caused another World Trade Center building to fall 7 hours later, and several

Continued on Page A14

Awaiting the Aftershocks

Washington and Nation Plunge Into Fight With Enemy Hard to Identify and Punish

By R. W. APPLE Jr.

News Analysis

WASHINGTON, Sept. 11 — Today's devastating and astonishingly well-coordinated attacks on the World Trade Center towers in New York and on the Pentagon outside of Washington plunged the nation into a warlike struggle against an enemy that will be hard to identify with certainty and hard to punish with precision.

The whole nation — to a degree the whole world — shook as hijacked airliners plunged into buildings that symbolize the financial and military might of the United States. The sense of security and self-confidence that Americans take as their birthright suffered a grievous blow, from which recovery will be slow. The aftershocks will be nearly as bad, as hundreds and possibly thousands of people discover that friends or relatives died awful, fiery deaths.

Scenes of chaos and destruction evocative of the nightmare world of Hieronymus Bosch, with smoke and debris blotting out the sun, were carried by television into homes and workplaces across the nation. Echoing Franklin D. Roosevelt's description of the attack on Pearl Harbor as an event "which will live in infamy," Gov. George E. Pataki of New York, a Republican, spoke of "an incredible outrage" and Senator Charles E. Schumer of New York, a Democrat, spoke of "a dastardly attack."

But mere words were inadequate vessels to contain the sense of shock and horror that people felt.

As Washington struggled to regain a sense of equilibrium, with warplanes and heavily armed helicopters crossing overhead, past and present national security officials earnestly debated the possibility of a Congressional declaration of war — but against precisely whom, and in what exact circumstances? Warships were maneuvering to protect New York and Washington. The North American Air Defense Command, which had seemed to many a relic of the cold war, adopted a pos

Continued on Page A24

MORE ON THE ATTACKS

RESCUERS BECOME VICTIMS Firefighters who rushed to the trade center were killed. PAGE A1

SEARCH FOR SURVIVORS Some people trapped in the rubble for hours were rescued. PAGE A1

OFFICIALS SUSPECT BIN LADEN Eavesdropping intercepts after the attacks were cited. PAGE A1

TERRORISTS EXPLOIT WEAKNESS Investigators had criticized precautions against hijacking. PAGE A1

CASUALTIES IN WASHINGTON An unknown number of people were killed at the Pentagon. PAGE A1

FOR HOME DELIVERY CALL 1-800-NYTIMES

Figure 1. The front page of the September 12, 2001 edition of *The New York Times*, showing two photographs in which the first World Trade Center tower to be struck by an airplane is on fire but still standing, moments before (smaller, right-hand image) and after (larger, centered image) the other tower was struck by a second airplane. (Copyright © 2001 *The New York Times*. Reprinted by permission.)

METHOD

Participants

Three groups of participants (453 in total) were recruited from the Michigan State University subject pool. There was one *November* group (158 participants), which was tested November 11 through 14, 2001 (roughly nine weeks after September 11). There were also two January groups, *January-six* (147 participants) and *January-five* (148 participants), which were both tested January 13 through 19, 2002 (roughly 18 weeks after September 11). The 'six' and 'five' suffixes distinguish the slightly different questionnaires administered to the two January groups, as described next. The three groups were disjoint, with no participants in common.

Materials

Table 1 shows the sets of events that participants were asked to sequence. Within a set, each event is identified (in the table) by its time of occurrence and its serial position within the set. On the actual questionnaire, time and serial position were not provided, and the order of events was randomized.

The six-event set consists of the crashes of the four hijacked airplanes plus the collapse of the two World Trade Center towers. This set was administered to the November and January-six groups.

The five-event set consists of the six-event set but without the crash of the airplane in Pennsylvania. The five-event set was developed after the November group had been tested, because in the resulting data the Pennsylvania crash was far less accurately sequenced than the other events (Figure 4). This low accuracy may reflect actual temporal and ordinal ambiguities. This crash occurred in a rural area, unlike the others, and there was a 20-minute lag before it was confirmed by local officials, such that confirmation of this crash came only after collapse of the second World Trade Center tower. Thus, both the time of this crash, and its serial position within the six-event set, were potentially ambiguous, so the five-event set was formed to remove these sources of variance. The five-event set was administered only to the January-five group.

The four-event set consists of President Bush's stopping points around the country that day. This set was administered to all three groups.

On the questionnaire, printed on a single 8.5×11 inch sheet of paper, the six- or five-event set appeared first (in the top half) and the four-event set appeared below it. Above each set of events was a short paragraph that provided some context and instructed participants how to sequence the events. In Table 1, the text of each paragraph appears in quotation marks.

Procedure

Participants were tested in sessions of 20 or fewer individuals. They took their seats at tables where the questionnaire had already been distributed. The session began with the experimenter simply asking participants to sign a consent form and then fill out the questionnaire. Participants were allowed as much time as necessary, but were typically finished in a few minutes, and were then debriefed with the actual timeline of all the events.

Table 1. Stimulus materials for testing serial-order reconstruction of events of September 11

Six-event set

‘On September 11, 2001, several airplanes were hijacked and crashed into locations in the northeastern United States. Six events from that day are listed below, in random order. Please number them in chronological order, with 1 being the first event to occur and 6 being the last.’

Time of event	*Events to be sequenced*
8:45 am	1. One plane hits the World Trade Center
9:03 am	2. A second plane hits the World Trade Center
9:43 am	3. One plane crashes into the Pentagon
10:05 am	4. One tower at the World Trade Center collapses
10:10/10:48 am[a]	5. One plane crashes in Pennsylvania
10:28 am	6. A second tower at the World Trade Center collapses

Five-event set

‘On September 11, 2001, several airplanes were hijacked and crashed into locations in the northeastern United States. Five events from that day are listed below, in random order. Please number them in chronological order, with 1 being the first event to occur and 5 being the last.’

Time of event	*Events to be sequenced*
8:45 am	1. One plane hits the World Trade Center
9:03 am	2. A second plane hits the World Trade Center
9:43 am	3. One plane crashes into the Pentagon
10:05 am	4. One tower at the World Trade Center collapses
10:28 am	5. A second tower at the World Trade Center collapses

Four-event set

‘That same day, President Bush was in four different locations around the country at various times. The locations are listed below, in random order. As best as you can remember, please number in chronological order, with 1 being his first location that day and 4 being his last.’

Arrival	*President’s address*	*Departure*	*Events to be sequenced*
	9:30 am	9:57 am	1. Sarasota, Florida
	1:04 pm	1:48 pm	2. Barksdale Air Force Base, Louisiana
	3:55 pm[b]	4:30 pm	3. Offutt Air Force Base, Nebraska
6:54 pm	8:30 pm		4. Washington, DC

Note: Events on the actual questionnaires were unnumbered and randomized, and the text in quotation marks was given to provide context and instruction.
[a]Time of occurrence/time of confirmation by local officials.
[b]Presidental advisor Karen Hughes announces that President Bush is at an undisclosed location.

Design and analysis

The independent variables were the serial position at which events occurred within their event set, and, for the four-event and six-event sets, retention interval (9 weeks vs. 18 weeks).

The dependent variables were the frequency of *correct responses*, in which events were placed in the serial position in which they occurred, and the full *response distribution*, including correct responses and errors.

The effects of serial position and retention interval on correct responses were examined with analysis of variance (ANOVA). The questions were whether a serial position curve would be evident and whether it would shift or change with retention interval.

The effect of serial position on the response distribution was examined by correlating each of the current distributions with one from an order-memory experiment involving the same serial-order reconstruction task and the same number of events per set. The question was whether response distributions would exhibit a positional gradient, with erroneous placements near to the correct placement (in terms of serial position) more frequent than those far away.

RESULTS

Data from the five-event set are plotted in Figure 2 (filled markers). In the top panel are correct responses—for example, the frequency with which event 1 was correctly placed in

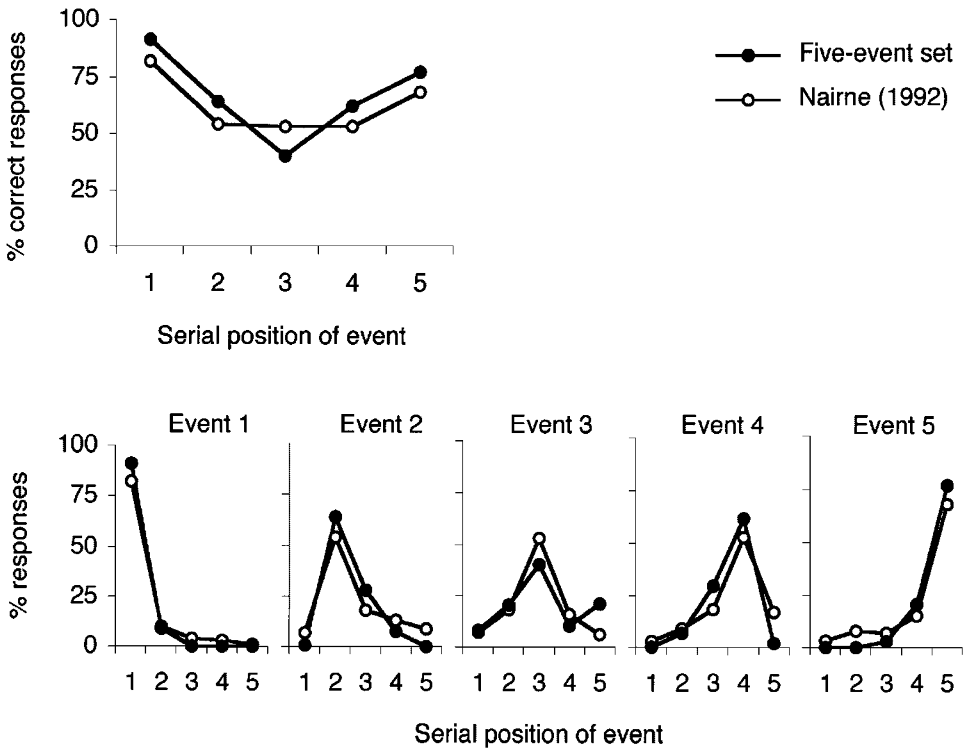

Figure 2. Correct responses (top) and all responses (bottom) in serial-order reconstruction of the five-event set (filled markers), compared to data from an order-memory experiment involving five-item lists (unfilled markers; Nairne, 1992)

position 1. To evaluate the serial position curve, correct responses were submitted to a one-way ANOVA, which showed a reliable effect of serial position, $F(4, 584) = 36.9$, $p < 0.001$.

The bottom panels of Figure 2 show full response distributions for the five-event set, including errors as well as correct responses. For example, the bottom left panel shows that 91.2% of participants placed the first of the five events correctly in position 1, that 8.8% placed that event incorrectly in position 2, and so on. For all but one event, response distributions follow a positional gradient in which the modal response is the correct one and error responses are more frequent the closer they are to the correct response. The one exception is event 3, for which the frequency of error increases from serial positions 4 to 5.

To assess the similarity of the positional gradient in the five-event set to that found in order-memory experiments, Figure 2 (unfilled markers) shows data from Nairne (1992). To recap, this study presented random words at a rate of 3 s each, with the nominal task to rate the 'pleasantness' of each word, then administered an incidental memory test 30 s later (in the condition shown in Figure 2). Despite the very different stimulus materials, presentation rates, and retention intervals, the Nairne data and the five-event data correlate 0.96 (RMSE = 7.8%).

Data from the four-event set, pooled across the three groups, are plotted in Figure 3 (filled markers). In the top panel are correct responses. To evaluate the serial position curve and the stability of performance over time, correct responses were submitted to a 4 × 2 ANOVA with factors serial position and test time (November vs. January-six pooled

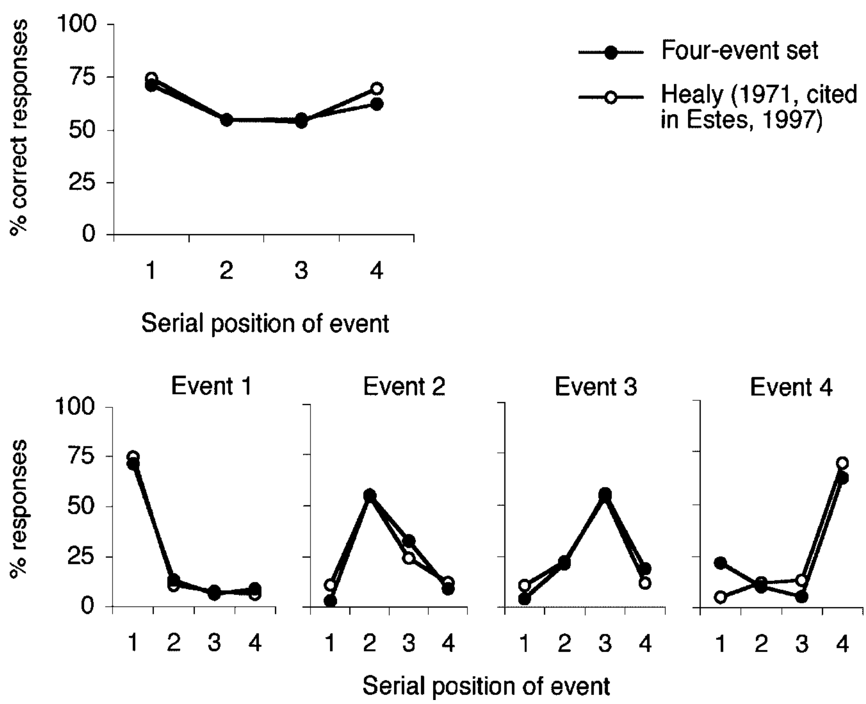

Figure 3. Correct responses (top) and all responses (bottom) in serial-order reconstruction of the four-event set (filled markers), compared to data from an order-memory experiment involving four-item lists (unfilled markers; Healy, 1971, as cited in Estes, 1997)

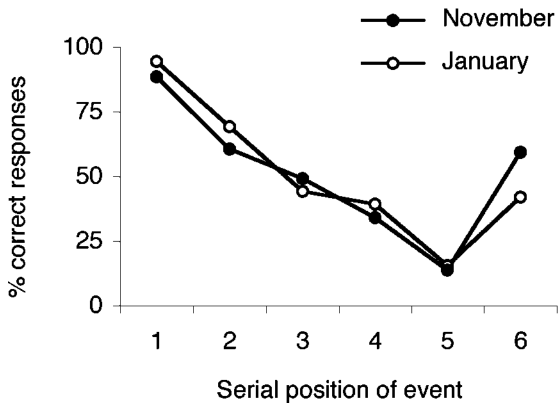

Figure 4. Correct responses in sequencing the six-event set in November (filled markers) and January (unfilled markers)

with January-five). There was an effect of serial position, $F(3, 1353) = 17.8, p < 0.001$, but no effect of test time, $F(1, 451) = 1.4$, $p = 0.24$, and no interaction, $F < 1$.

The bottom panels of Figure 3 show full response distributions for the four-event set. For all but one event, response distributions again follow a positional gradient. The exception is event 4, for which the frequency of error increases from serial positions 3 through 1.

To assess the similarity of the positional gradient of the four-event set to that found in order-memory experiments, Figure 3 (unfilled markers) shows data from Healy (1971, as cited in Estes, 1997), who tested intentional memory for the order of four random consonants after filled retention intervals ranging from 6 to 36 s. The four-event data and the Healy data correlate 0.96 (RMSE = 6.4%).

Finally, data from the six-event set are plotted in Figure 4, separated by test time. Submitted to a 6×2 ANOVA on serial position and test time (the November group vs. the January-six group), correct responses showed an effect of serial position, $F(5, 1520) = 124.0$, $p < 0.001$, but not of test time, $F \approx 0$, though the two did interact, $F(5, 1520) = 4.4, p < 0.002$. Because of the very low accuracy on event 5, which may have reflected a temporal and ordinal ambiguity (as discussed above), no attempt was made to compare the six-event set to data from order-memory experiments.

DISCUSSION

The primary empirical question was whether patterns of serial-order reconstruction of events of September 11 would mirror those from order memory experiments, or whether the very different circumstances of storage and encoding would translate into qualitatively different response distributions. Figures 2 and 3 suggest that standard order-memory experiments fare rather well as models of the current data. Indeed, the correlations (0.96 in both figures) rank at the top of the range achieved by theoretical models designed specifically to account for the Nairne data (0.93 to 0.96; Anderson & Matessa, 1997). The general implication is that across a wide range of circumstances, serial-order reconstruction will exhibit the positional gradient and serial position curve (a point revisited again below).

As a case study of serial-order reconstruction of real-world events, the current findings suggest that even when people are able to reconstruct the 'gist' of a sequence, placing most events correctly most of the time, they often get neighboring events out of order. On the five-event set, 63% of participants made at least one error, and of these, 72% fell within one position of the correct response. On the four-event set, the statistics were similar; 59% of participants made at least one error, and of these errors 62% fell within one position of the correct response. These high error rates may offer yet another reason to be skeptical of eyewitness accounts of critical event sequences. Quite apart from failures of declarative memory usually studied in research on eyewitness memory (e.g., Loftus, 1996; Yuille & Cutshall, 1986), serial-order reconstruction can distort in a way that could materially affect causal reasoning by reversing the order of two neighboring events—even events that the observer may consider 'unforgettable.'

On the positive side, serial-order reconstruction was stable over time, with no effect of doubling the retention interval from 9 weeks to 18 weeks. This null effect of retention time suggests that frequent rehearsal, internally or through conversations or exposure to popular media, serves to stabilize the underlying memory representations. Rehearsal has been found to stabilize other autobiographical memories as well (e.g., Neisser et al., 1996). However, it may also contribute to distortion (e.g., Schmolk, Buffalo, & Squire, 2000; Talarico & Rubin, 2003), and at least one memory theory specifies rehearsal as a source of distortion (Estes, 1997), so there are clearly other contingencies involved. The current data say nothing about changes in performance over the initial 9 weeks after September 11, during which most of any forgetting may have taken place (Burt, Watt, Mitchell, & Conway, 1998; Winningham, Hyman, & Dinnel, 2000).

The primary theoretical question was how to explain a positional gradient in situations where ordinal information may come more from relationships between events in a narrative than from direct experience of the events in the order in which they occurred. To address this question, Monte Carlo simulations were used to explore whether a specific set of simple constraints is in principle sufficient to produce the empirical patterns in the September 11 data. The focus is on the five-event set, because several pairwise constraints were inherent in the stimulus materials. These *logical* and *lexical* constraints are formalized in Table 2. Logically, for example, neither World Trade Center tower would have collapsed had it not been struck first by an airplane. These are similar to the constraints that bind memories of goal-related activities (e.g., Bower, Black, & Turner, 1979; Lichtenstein & Brewer, 1980), and to logical constraints that developmental studies suggest are quite basic to reasoning about event sequences (Bauer, Hertsgaard, Dropik, & Daly, 1998; Fivush, Kuebli, & Clubb, 1992; Hudson & Nelson, 1983). Lexically, certain constraints were difficult to avoid in formulating the stimulus materials, given the nature of the events. The two towers of the World Trade Center and the two airplanes that flew into them had to be differentiated in the stimulus materials, without assuming unrealistically detailed knowledge on the part of participants about New York (the addresses or relative locations of the towers) or the airplanes (their flight numbers or other unique attributes). The solution adopted here was to differentiate the towers and the airplanes in the stimulus materials by their ordinal position—hence the use of 'One plane' and 'A second plane,' and 'One tower' and 'A second tower,' in the event descriptions in Table 1.

To examine how these lexical and logical constraints might have affected serial-order reconstruction, random orderings of five events were generated and then filtered to remove orderings that violated the constraints. The resulting simulated response distributions are shown in Figure 5, compared with the empirical data from the five-event set. The panels

Table 2. Constraints on the serial order of the five-event set

Type of constraint	Pairwise constraint		
Linguistic	One plane hits	*precedes*	A second plane hits
	One tower collapses	*precedes*	A second tower collapses
Logical	One plane hits	*precedes*	One tower collapses
	One plane hits	*precedes*	A second tower collapses
	A second plane hits	*precedes*	A second tower collapses
Grouping	One plane hits	*precedes*	One tower collapses[a]
	One plane hits	*precedes*	A second tower collapses[a]
	A second plane hits	*precedes*	One tower collapses
	A second plane hits	*precedes*	A second tower collapses[a]

Note. [a]Redundant with a logical constraint.

across the top show the effect of the lexical constraints, which would seem to be the most difficult to avoid applying at test time. The mismatch, apparent visually and supported statistically by a χ^2 test[1] indicates that these constraints are insufficient to reproduce the empirical data. The middle panels show the effects of adding the logical constraints; the fit is now closer, particularly for the first and last events, but fails to capture the modal correct placement of the middle three events, and the χ^2 test again indicates a reliable mismatch.

This exhausts the constraints inherent in the stimulus materials, so we are now left to guess at what additional constraints stored in memory might have been implicated. One possibility is that participants remembered that both World Trade Center towers were hit before either collapsed. Among the most common images of September 11 are those of the moments immediately preceding and following the crash of the second airplane into the World Trade Center, and in many such images (like the two in Figure 1) both towers are visible, as a function of their proximity. When the *grouping* constraint (Table 2) implied by such images is added to the lexical and logical constraints, the simulation produces the response distribution in the bottom panels of Figure 5. The fit here is reasonably good, with the simulation correlating 0.96 (RMSE = 7.0%) with the empirical data, and the χ^2 test failing to detect a mismatch. Qualitatively, the simulation predicts that event 3 (the crash into the Pentagon) should have been placed at chance, whereas empirically the modal response was correct for this event as well. This mismatch could reflect the influence of an ordinal-context representation, acquired by actually experiencing reports of the Pentagon crash in order relative to the other events of the five-event set. Apart from this mismatch, however, this handful of constraints—some inherent in the stimulus materials, one hypothetically coded in memory—explains much of the variance in participants' performance.

In situations like the four-event set, where no similar constraints are inherent in the stimulus materials, any account of the positional gradient is necessarily more speculative. An oft-replayed image was that of President Bush sitting in a Florida classroom looking pensive as he received news of the first collision into the World Trade Center. With Florida

[1]The χ^2 statistics in Figure 5 were computed on the 5×5 matrices produced by crossing position in the original sequence with position placed at test. However, to prevent the statistic from being driven by relatively unstable values, a cell was excluded if either its expected (simulated) or empirical value contained fewer than 10 observations (in the limit, a cell with no observations would make the statistic undefined). Degrees of freedom were computed by first subtracting five from the number of cells of the matrix (25), because for each item, the proportion of placement in four positions determines the proportion placement in the fifth, and then by subtracting the number of cells with fewer than ten values.

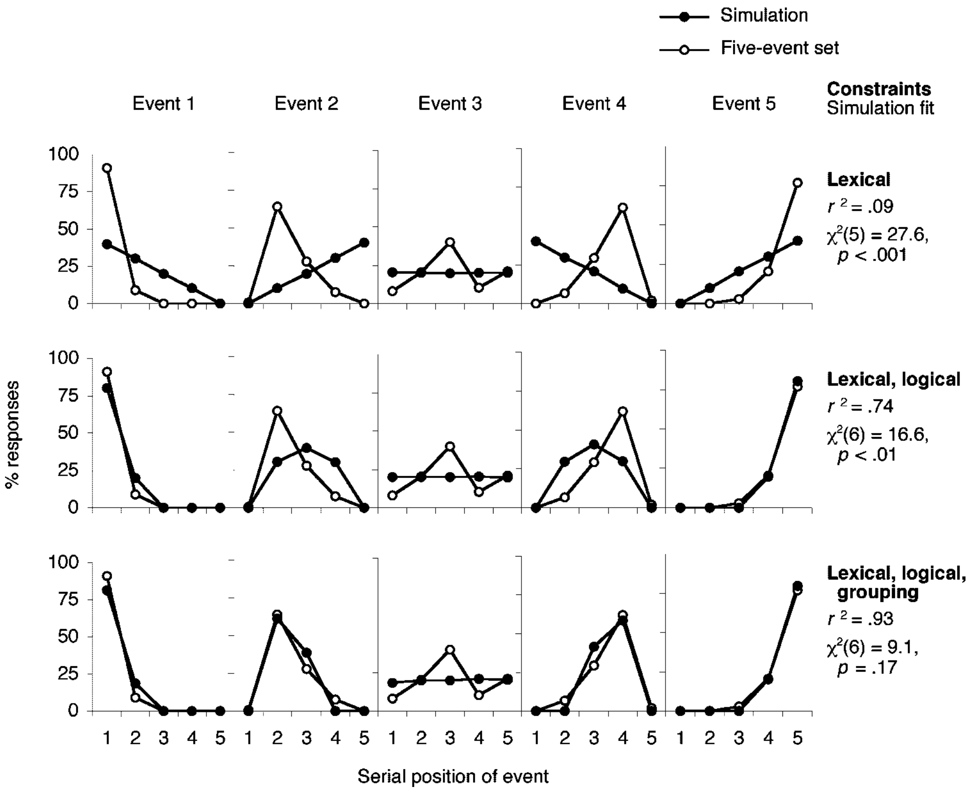

Figure 5. Responses in sequencing the five-event set (unfilled markers) compared to data from Monte Carlo simulations (filled markers) incorporating the constraints described in Table 2

anchoring the start of this sequence, there is a certain logic to Washington being last, and geographical constraints suggest that Louisiana was more likely first of the two remaining stops. Alternatively, an ordinal-context representation may explain the positional gradient, given that these four events were distributed throughout the day and thus may well have been experienced by many observers in the order in which they occurred. A third alternative is an explanation in terms of simple *order codes* relating pairs of events learned associatively when one event primes retrieval of a related event (Friedman, 1993). Evidence for such codes comes from the event-cuing method of probing autobiographical memory (N. R. Brown & Schopflocher, 1998), and given the pervasiveness of associative learning, such order codes might be a widely-available fallback representation when more specific recollections and reconstructive logic are not available.

This study indicates a number of directions for future research on serial-order reconstruction, involving both real-world events and events presented in more controlled circumstances. In the laboratory, studies with ordinal information coded descriptively rather than experienced directly may show that the two representations are interchangeable—or conceivably that ordinal-context representations are not required. For example, the pleasantness-rating task often used when order memory is tested incidentally (e.g., Nairne, 1992) could evoke images for purposes of rating pleasantness that then bind neighboring events associatively, much like the images in Figure 1 bind events of the five-event set.

In terms of real-world events, the various possible accounts of the four-event data discussed above illustrate that when events are related or embedded in a narrative structure of some kind, there could be any variety of explanations for the positional gradient. One hypothesis, then, is that the positional gradient is a product of the nature of the serial-order reconstruction task itself, rather than any particular set of constraints. Thus, the task may trigger reconstructive processes that draw on whatever constraints are available—ordinal-context representations acquired through experience, semantic constraints inferred from descriptions, or order codes acquired through associative learning—to place each event as close to its original ordinal position as possible, producing the positional gradient as a side effect. Evidence for this hypothesis would have to accrue through broad testing of serial-order reconstruction in many different circumstances, to show that the positional gradient is in fact the common denominator.

ACKNOWLEDGEMENTS

This research was supported in part by a grant from the Office of Naval Research. Thanks to Bruce Burns, Zach Hambrick, and Kathy Swedlow for comments on the draft, to Kara Devlieger for help in data collection, and to two anonymous reviewers for their thoughtful suggestions for improvement.

REFERENCES

Anderson, J. R., & Matessa, M. (1997). A production system theory of serial memory. *Psychological Review, 104*, 728–748.

Bauer, P. J., Hertsgaard, L. A., Dropik, P., & Daly, B. P. (1998). When even arbitrary order becomes important: developments in reliable temporal sequencing of arbitrarily ordered events. *Memory, 6*, 165–198.

Bower, G. H., Black, J. B., & Turner, T. J. (1979). Scripts in memory for text. *Cognitive Psychology, 11*, 177–220.

Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review, 107*, 127–181.

Brown, N. R., & Schopflocher, D. (1998). Event clusters: an organization of personal events in autobiographical memory. *Psychological Science, 9*, 470–475.

Burt, C. D. B., Watt, S. C., Mitchell, D. A., & Conway, M. A. (1998). Retrieving the sequences of autobiographical event components. *Applied Cognitive Psychology, 12*, 321–338.

Estes, W. K. (1997). Processes of memory loss, recovery, and distortion. *Psychological Review, 104*, 148–169.

Fivush, R., Kuebli, J., & Clubb, P. A. (1992). The structure of events and event representations. *Child Development, 63*, 188–201.

Friedman, W. J. (1993). Memory for the time of past events. *Psychological Bulletin, 113*, 44–66.

Hudson, J., & Nelson, K. (1983). Effects of script structure on children's story recall. *Developmental Psychology, 19*, 625–635.

Huttenlocher, J., Hedges, L. V., & Bradburn, N. M. (1990). Reports of elapsed time: bounding and rounding processes in estimation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 196–213.

Huttenlocher, J., Hedges, L. V., & Prohaska, V. (1992). Memory for day of the week: A $5 + 2$ day cycle. *Journal of Experimental Psychology: General, 121*, 313–325.

Lichtenstein, E. H., & Brewer, W. F. (1980). Memory for goal-directed events. *Cognitive Psychology, 12*, 412–445.

Loftus, E. F. (1996). *Eyewitness testimony*. Cambridge, MA: Harvard University Press.

Loftus, E. F., & Marburger, W. (1983). Since the eruption of Mt. St. Helens, has anyone beaten you up? Improving the accuracy of retrospective reports with landmark events. *Memory & Cognition, 11*, 114–120.

Nairne, J. S. (1992). The loss of positional certainty in long-term memory. *Psychological Science, 3*, 199–202.

Neisser, U., Winograd, E., Bergman, E. T., Schreiber, C. A., Palmer, S. E., & Weldon, M. S. (1996). Remembering the earthquake: direct experience vs. hearing the news. *Memory, 4*, 337–357.

NTSB. (2002). *Aircraft accident brief: EgyptAir 900* (No. NTSB/AAB-02-01, PB2002-910401). Washington DC: National Transportation Safety Board. Retrieved from http://www.ntsb.gov/Publictn/2002/aab0201.htm.

Page, M. P. A., & Norris, D. (1998). The primacy model: a new model of immediate serial recall. *Psychological Review, 105*, 761–781.

Schmolk, H., Buffalo, E. A., & Squire, L. R. (2000). Memory distortions develop over time: recollections of the O. J. Simpson trial verdict after 15 and 32 months. *Psychological Science, 11*, 39–45.

Shum, M. S. (1998). The role of temporal landmarks in autobiographical memory. *Psychological Bulletin, 124*, 423–442.

Talarico, J. M., & Rubin, D. C. (2003). Confidence, not consistency, characterizes flashbulb memories. *Psychological Science, 14*, 455–461.

Wagenaar, W. A. (1986). My memory: a study of autobiographical memory over six years. *Cognitive Psychology, 18*, 225–252.

Winningham, R. G., Hyman, I. E., Jr., & Dinnel, D. L. (2000). Flashbulb memories? The effects of when the initial memory report was obtained. *Memory, 8*, 209–216.

Yuille, J. C., & Cutshall, J. L. (1986). A case study of eyewitness memory of a crime. *Journal of Applied Psychology, 71*, 291–301.

