

Despite the survey design elements used to prompt recall, we were concerned about respondents' ability to remember, and remember accurately, events from ODS/DS. There is no way to conclusively evaluate how well our survey population recalled exposure to pesticides without recourse to detailed exposure records, which in our case do not exist. Therefore, we cannot know definitively whether some survey respondents were systematically underreporting exposure or other respondents were overreporting exposure. Both add error to our measure of exposure; in the end they may neutralize each other.

However, we felt compelled to try to evaluate how recall might affect our results. In the absence of such an evaluation of the accuracy of recall, we would have had to assume that respondents' responses reflected exactly what occurred during ODS/DS. However, we know that recall of events almost a decade in the past is likely to be imperfect. This chapter examines the extent of this imperfection by comparing the follow-up (recall bias) survey we administered to the main survey data.

#### **METHODS FOR ASSESSING RECALL BIAS**

We reviewed the scientific literature on recall bias and learned that memory can be unreliable in two ways. First, some details of an experience may never be noticed or stored in memory. For example, personnel may not be aware of the pesticides used in their mess halls. Second, information may be added later if memories are "rehearsed," that is, events are recalled by thinking or talking about them and then re-stored in memory. Rehearsal increases the ease with which we can recall memories, and failure to rehearse or recall a memory for a long time can make it difficult or impossible to retrieve it when it is wanted. However, rehearsal can also contaminate the original memory: When the memory of an event is recalled to consciousness, other new "facts" about the event may be added as the event is embellished, made more socially acceptable, redefined to fit present-day conceptions, or appended in any number of

additional ways. When the memory is again stored in long-term memory, it may be stored in an altered fashion that includes new information. If the altered memory is the one that is most rehearsed, then it is likely to become the perceived “real” memory.

We were less concerned with this aspect of recall, as pesticide exposure has not been prominent among Gulf War issues. What it does highlight, however, are the two aspects to remembering: ability and effort. We reviewed the extensive literature on questionnaire design, memory, and recall to avoid where possible the methodological pitfalls to which self-reports of exposure may be prone, and we designed the main survey with the findings from this literature in mind. For example, the literature suggests that easily demarcated events—such as a war—are easily recalled, but mundane day-to-day events—such as pesticide use—may not be. Our goal was to construct a survey that would aid accurate recall by helping respondents reconstruct the context of their experiences. This included questions regarding attributes of respondents’ living and working environments, questions on the kinds of pests they faced, and questions designed to help them reconstruct a timeline of their experiences. These questions were intended to encourage their recall of the day-to-day aspects of their life while in the Gulf region.

### **Assessing Recall Bias Through Re-Survey**

To evaluate the effect of recall on our survey results, we administered a second survey to a small sample of initial survey respondents. To avoid overinterpreting any one question or type of question, we surveyed and examined multiple dimensions along which systematic recall bias may have occurred, such as service, pay grade, education, and reported health status. In short, we employed multiple tests of recall bias in the knowledge that no single measure could accurately capture the extent of recall bias as a whole. Multiple tests avoid overinterpreting any one measure, question, or set of questions. This is also important because there may be offsetting biases, none of which could be predicted in advance of the survey. Some analysts find that recall bias contaminates their results; other analysts conclude that recall bias does not affect respondents’ answers. Our purpose in administering the recall bias survey and otherwise assessing the reliability and consistency of respondents’ answers is to place an honest range of uncertainty around our estimates of pesticide exposure. This allowed us to assess the level and direction of bias, and to clearly report the sensitivity of our results to it.

We cannot know definitively how well survey respondents recall pesticide exposure. However, the literature suggests that we can gauge the extent to which recall might affect our results by examining how changes in reported use at re-

survey vary by certain individual characteristics. These include health status and sensitivity to the issues of pesticide use and Gulf War illnesses, as well as other demographic factors such as education.

### **Recall Bias and Health Status**

We employed a commonly asked question designed to elicit information about the respondent's current health. As discussed above, the ability to recall past events is in part dictated by people's willingness to put the effort into remembering. Currently ill respondents already will have invested time and energy into thinking about their health and may be both more sensitized to public discussions of the issue and more attentive to factors hypothesized to negatively affect health. This can cause them to overreport exposure if in the process of remembering they have assimilated the experiences of others into their own. Conversely, illness may cause underreporting of exposure, if current illness hampers respondents' ability to concentrate, for example. Thus, it is not possible to predict in advance which effect will dominate, but collecting data on current health status helps to determine whether a possible problem with recall bias exists.

To reiterate, we are not able to draw firm conclusions about exposure and recall bias solely using information on health status. If ill respondents report more exposure, for example, this could be interpreted several ways: (1) It could be true; (2) it could be because they have been following the debate and talking to others about their experiences; or (3) it could be that healthy respondents are not interested enough in this issue and therefore do not put the same level of effort into remembering. Our objective in collecting the data is simply to document whether responses vary by current health status, rather than to draw definitive conclusions.

### **Recall and Sensitivity to Gulf War Issues**

Further, the public controversy over Gulf War illnesses could affect how much effort the respondent puts into recalling pesticide use. Respondents uninterested in revisiting issues related to the war may try to rush through the survey; respondents following such issues more closely may take more time to try to remember. We assessed these potential markers of recall bias in several ways. One method we used was to ask early in the survey (before trying to elicit memories of exposure) how much interest the respondents have generally had in Gulf War issues and whether they have thought much about the pesticides they encountered during their tours of duty. Another method was to ask whether respondents reporting fair or poor current health thought their health status was linked to the Gulf War or whether their doctors thought so. We also asked re-

spondents if they had registered with the Veteran's Administration or Department of Defense registries, since veterans who have registered may report higher or lower levels of pesticide exposure. Although there is no causal conclusion to be drawn from such an association, we wanted to reveal any systematic patterns of differential response within our sample. One important dimension along which they might differ would be the extent to which they are presently engaged in Gulf War issues.

## RESULTS

The recall bias survey sample is approximately 8 percent of the full sample, which means that it is large enough to statistically detect changes in answers that most survey respondents gave, but not in some of the less common answers. Thus, our initial analysis focuses on stability in aggregate measures: number of pests and types of pests observed, number of types of personal pesticides used, and number of types of field pesticides used. We also concentrated more on the patterns across subgroups and across outcomes than on statistical significance. We then examined personal-use sprays and animal traps, the most commonly reported personal and field-use pesticides, in more detail. This part of the analysis included how they were used, and whether the same pesticide name was reported in both surveys.

We found evidence of changes in responses overall, with the fraction reporting pesticide types increasing about 13 percent in the re-survey. We did not see strong patterns among the various groups in our data; this includes not only demographic groupings, such as education or rank, but also self-reported health status. However, we also found that people who thought about their pesticide exposure before our survey reported more pesticide use, but their answers were in fact more stable over time. We interpreted the pattern of differences as an indication that people who had not thought about pests and pesticides since the war were less likely to put as much effort into recalling their experiences for our survey. Answers on how pesticides were used (such as number of sprays used or frequency of use) were stable across surveys. In general, most respondents did not report names of pesticides in either survey, nor did they use most types asked about. This remains the most salient finding across the two surveys.

### Analysis of Recall Bias Across Groups

To compare answers across various groups in the data, such as service or rank, we examined responses to questions about aggregate pesticide usage, such as

number of types of pesticides used or observed. We also evaluated the number of kinds of pests reported.<sup>1</sup> The results are reported in Table D.1. Overall, the 193 respondents in the follow-up sample reported seeing 5.28 types of pests on average in the original survey; in the follow-up, they reported 0.22 fewer types, or 5.06 types of pests. The 4 percent drop is statistically significant. They also reported using 0.88 types of personal-use pesticides (liquids, sprays, powders, etc.) and reported 0.99 types in the follow-up survey, a statistically significant increase of 13 percent. The number of types of field-use pesticides reported in the follow-up survey also increased by 13 percent and, again, this was statistically significant. As people had more time to consider their answers during the period between surveys, they might have simply convinced themselves that they saw more or experienced more of everything. Instead, the results suggest that recall of *pests* was more reliable than recall of *pesticides*. It appears that respondents considered their experiences more thoroughly in light of our questions and answered more carefully the second time.

Aware that, given time, respondent' answers could change, we looked for evidence of systematic bias in their answers. We examined multiple dimensions along which we might expect to see such bias, such as service, pay grade, education, and reported health status. Although the total sample size is 193, dividing the sample to look for subgroup differences is statistically difficult, as the power of the tests is reduced due to small numbers. Thus, we did not necessarily expect to find statistically significant results in this part of the analysis. Instead, we placed more weight on the patterns across subgroups and across outcomes, and there we interpreted the results as lacking evidence of systematic bias among subgroups in the survey sample.

The results of this analysis are shown in Table D.1. Although Army members reported many fewer pest types, they did not exhibit the most change in personal-use pesticides (the Air Force did, percentage-wise) nor of field-use types (the Marine Corps showed the largest percentage change). Similarly, African-American veterans report the largest change in the number of pests, Caucasians the largest change in number of personal-use pesticides, and other races the largest changes in field-use types. More educated respondents remembered relatively more personal-use types in the follow-up survey, whereas less educated respondents remembered relatively more field-use applications. The only group whose answers changed in statistically significant ways for all three variables were junior enlisted personnel (pay grades E-1 to E-5), who remembered fewer pests and more pesticides, both personal and field-

---

<sup>1</sup>This question was used to prompt memories of pesticide use by encouraging respondents to recall why they needed pesticides. We did not expect respondents' answers to change across surveys, and so we use this as a gauge of the magnitude of the change in the pesticide measures.

**Table D.1**  
**Correlates of Recall**

	Average Number of Pests	Change in Average Between Surveys	Average Number of Personal-Use Pesticides	Change in Average Between Surveys	Average Number of Field- Use Types	Change in Average Between Surveys
Overall	5.28	-0.22*	0.88	0.11*	1.35	0.18**
Air Force	5.45	-0.04	0.66	0.14*	1.50	0.16
Marine Corps	5.23	-0.14	1.15	0.02	1.38	0.29**
Army	5.45	-0.59**	0.89	0.14	1.20	0.07
Navy	++	++	++	++	++	++
Caucasian	5.15	-0.08	0.83	0.13**	1.30	0.18**
African-American	5.86	-0.76*	1.07	0.03	1.52	0.14
Other	5.38	-0.43	0.95	0.05	1.48	0.24
Male	5.37	-0.21	0.91	0.11**	1.41	0.22**
Female	4.44	-0.33	0.50	0.06	0.83	-0.17
Active	5.33	0.03	0.78	0.23**	1.63	0.20
Reserves	4.76	-0.04	1.00	0.08	1.04	0.32*
Retired	5.09	-0.14	0.70	0.05	1.32	0.09
Civilian	5.51	-0.43**	0.98	0.10	1.33	0.18*
E-1 to E-5	5.52	-0.32*	0.92	0.12**	1.33	0.20**
E-6 to E-9	5.20	-0.15	0.7	0.08	1.45	0.03
Officer	4.11	0.16	1.00	0.11	1.16	0.53*
High school or less	5.08	-0.21	0.85	0.03	1.39	0.21
Some college	5.79	-0.13	0.86	0.18**	1.35	0.15
College graduate	4.61	-0.42	0.95	0.13*	1.29	0.18

NOTE: The sample used for estimation was the follow-up sample (n = 193).

\*\*p-value  $\leq$  0.05; \*p-value  $\leq$  0.1 from paired t-test of original average to follow-up average.

++Answer suppressed because there are under 10 cases in the cell.

use. This is somewhat, but not entirely, correlated with age, as younger respondents were more likely to be junior enlisted.

One unusual result we found—likely related to recall bias but not related to the recall survey—was that personnel currently on active duty tended to give names of military pesticides whereas civilians tended to give names of nonmilitary pesticides. That is, named pesticides tended to be related to a respondent's current status. We attribute this differential to recall, with current active duty personnel likely having been more recently aware of military products. In contrast, civilians are less likely to have recently been in contact with military products and more likely to have used or otherwise been in contact with nonmilitary products.

## Reported Health and Awareness of the Research Hypothesis

We were interested in how perceived health affects responses. For example, we were concerned that poor health might give respondents an extra incentive to think about their experiences and report pesticide use, or that poor health might inhibit memory. However, we also knew that self-reported health measures may not be reliable indicators of actual health and may be influenced by question wording, in particular, the order in which the responses are presented (Means et al., 1989). Thus, we randomly assigned to half of the original sample a question that asked to them rate their health from excellent to poor; the other half of the sample were asked to rate their health from poor to excellent.

As expected, we found that when excellent was the first response presented, as shown in Table D.2, respondents reported better health on average than when poor was the first response presented: 47 percent replied that their health was excellent or very good when those answers were presented first, compared with 36 percent of the other group. The difference is statistically significant ( $p = 0.07$ ).

Nonetheless, other survey responses appear to be relatively unaffected by health status and by which version of the question was asked. Respondents in both fair/poor and very good health reported seeing more pests in the initial survey than did respondents in good or excellent health, and those reporting fair/poor health also reported fewer pests in the follow-up survey. We cannot explain this odd pattern, having expected to see a smoother change across categories, and so we interpret this to mean that there is no systematic bias by health status. More important, there were no significant differences by health status or question version for number of personal pesticides used and number of field applications witnessed. This is shown in Table D.3, which reports the

**Table D.2**  
**Self-Reported Health Status (percent)**

Health Status	Version A (n = 86)	Version B (n = 107)	Version C (n = 193)
Poor	5.8	2.8	4.2
Fair	10.5	23.4	17.6
Good	37.2	38.3	37.8
Very good	26.7	25.2	25.9
Excellent	19.8	10.3	14.5

NOTES; Response categories were read aloud to the survey respondent. Version A of the question ordered response categories from excellent to poor; Version B was ordered poor to excellent. The sample used in estimation was the follow-up sample (N=193).

Table D.3

**Coefficients from a Regression of Levels and Changes Between Surveys on Health Measures and Question Wording**

Health Status	Average Number of Pests	Change in Average Between Surveys	Average Number of Personal-Use Types	Change in Average Between Surveys	Average Number of Field-Use Types	Change in Average Between Surveys
Poor/fair	1.23**	-0.77*	0.16	-0.08	0.10	-0.03
Good	1.00	-0.34	0.31	0.01	-0.02	0.04
Very good	1.06*	-0.59	0.18	-0.03	-0.05	0.09
Version A	0.20*	-0.12	0.01	0.07	0.03	-0.04
Constant	4.27**	0.29	0.67**	0.10	1.34**	0.17
R-squared	0.03	0.02	0.01	0.01	0.00	0.00

NOTE: The sample used for estimation was the follow-up sample (n = 193).

\*\*p-value  $\leq$  0.05, \*p-value  $\leq$  0.1.

coefficients from a regression of number of pests (number of types of use) on the health measures and the version of the question asked.

We also explored how these estimates changed when we included whether someone in poor/fair health had reported being enrolled in a Gulf War Registry. For the most part, the estimates remained similar to those reported above. It was interesting that registrants remembered more pesticides (both personal and field-use), and their answers across surveys were more stable regarding the number of types of personal pesticides they used. This is in keeping with the initial survey's questions about how much respondents had thought about pests and pesticides, with those answering "a lot" reporting more pests and more personal pesticide use; additionally, their answers did not change as much across surveys. As shown in Table D.4, those who reported in the initial survey that they had thought very little ("almost none") about pests and pesticides before the interview (most of the sample—see the Introduction) also reported fewer pesticides types in the second interview. The survey asked difficult-to-remember questions about events eight years before the interview. We suspect that respondents who had not thought about pests and pesticides in the intervening years did not put as much effort into remembering their experiences the first time through the survey as did the rest of the sample.

### Change Across Surveys in Pesticide Use

In both the original and follow-up surveys, we asked whether a particular spray, lotion, or other personal pesticide was used on the body, on the uniform, or both. These answers did not change much. Spray use was most likely to change, and in a pattern we did not anticipate: 10 percent changed their



**Table D.4**  
**Awareness of Gulf War Issues**

	Average Number of Pests	Change in Average Between Surveys	Average Number of Personal- Use Types	Change in Average Between Surveys	Average Number of Field-Use Types	Change in Average Between Surveys
Overall	5.28	-0.22*	0.88	0.11**	1.35	0.18**
Before today, how much have you thought about your Gulf War experiences in general?						
A lot	5.67	0.42*	0.97	0.00	1.30	-0.07
Some or a little	5.21	0.06	0.89	-0.17**	1.36	-0.28**
Almost none	4.27	0.67	0.40	-0.07	1.53	0.13
Before today, how much have you thought about problems you had with pests, rats, or other pests in the Persian Gulf, and the pesticides you used to get rid of these problems?						
A lot	6.00	0.65	1.18	-0.29*	1.53	-0.18
Some or a little	5.78	0.08	1.12	-0.06	1.49	-0.16*
Almost none	4.76	0.25	0.64	-0.11**	1.21	-0.20**

NOTE: The sample used for estimation was the follow-up sample (n = 193).

\*\*p-value  $\leq$  0.05; \*p-value  $\leq$  0.1 from t-test of original average to follow-up average.

answer from both to just one type of use. Nonetheless, 86 percent reported the same answer in both surveys.

The results for whether someone named the pesticide in either or both surveys were similar. People who named pesticides in the first survey named fewer pesticides in the second survey. We did not expect that. Yet very few named a spray (the most common personal-use pesticide form) in either survey. Of the 24 who provided a name in the follow-up survey, 96 percent gave the same name. It is easy to lose sight of the fact that 83 percent did not change the number of names they provided across surveys, whether they specified a name or not. In other words, few people remembered pesticides by name and this did not change substantially across the two surveys.

We also asked about the number of personal pesticides respondents used by type and the number of field applications they observed. Personal use appears to be stable when taken as a whole—74 percent reported using exactly the same number of sprays in both surveys, 12 percent reported more sprays, and 12 percent reported fewer sprays.<sup>2</sup> Reported frequency of use also remained stable across surveys. As shown in Table D.5, one interpretation of this result is that field use could be underestimated on average in the main survey. The extent of the difference was statistically significant at the 95 percent significance level or better for traps, pellets, and sprays from trucks. Again we note,

<sup>2</sup>Ninety-four percent of those who report using a liquid give identical answers across surveys about the number of liquids. The other forms do not have at least 10 people reporting use, and so we do not analyze the answers.

**Table D.5**  
**Percentage Reporting Field Use of Pesticides Across Surveys**

	Reported in Original Survey, But Not in Follow-Up	Answer Did Not Change	Did Not Report in Original Survey But Did in Follow-Up	Change in Av- erage Percent- age Reporting Use
Animal traps	6.5	80.5	13.0	+12**
Powders	3.9	89.5	6.6	+30
Pellets	0.6	95.5	3.9	+25**
Aerosol	7.0	85.0	8.0	+4
Spray from a truck	1.6	93.6	4.8	+11**
No-Pest strips	3.2	91.4	5.4	+44

NOTE: We examined only forms for which at least 10 people in the follow-up survey reported observing field use.

\*\*p-value  $\leq 0.05$ .

however, that most answers did not change, largely because most people reported no field use in either survey.

## CONCLUSION

The frequency of reporting of pesticide types increased in the re-survey by 13 percent. This change occurred primarily among respondents who had given less thought to their Gulf War experiences in the intervening years, but was not systematically related to other individual characteristics. We hesitate to provide a specific interpretation of these results given the complex nature of recall bias and the fact that ultimately we are only measuring differences between the two surveys—differences that could occur for many reasons. However, a “worst case” interpretation of the results is that the incidence of pesticide reporting could be underestimated in the initial survey.

Although the overall frequency of pesticide use may be somewhat higher than the survey results show, there is no evidence that different pesticides were subject to different levels of recall bias. This was qualitatively true even for field use, which showed varying degrees of change according to the type of application reported. There we found somewhat large percentage changes but low overall reporting, and we note that it is easy to lose track of the fact that a large percentage increase in a small number is still a small number. Therefore, we conclude that the mix of pesticides reported in the main survey does not appear to be misestimated.