Evidence of Potential Fabrication in Iraq Public Opinion Polls from 2005 – 2008

Michael Spagat

Department of Economics

Royal Holloway, University of London

August, 2011

Abstract

There have been many surveys of Iraqi public opinion since the onset of war in 2003. However, the underlying datasets have not been adequately scrutinized to ensure their authenticity. In this paper we provide evidence of potential data fabrication in the Iraq polling operations of a major research fielding agency. We examine datasets from five polls conducted by this contractor between 2005 and 2008 and find patterns in the data that suggest fabrication among a specific group of supervisors working in Baghdad, Anbar, and Diyala provinces. This fielding agency worked with a wide variety of private and U.S. Government clients during the period in question. As such, if our suspicions are confirmed, it casts doubt on a large body of accumulated research on Iraqi public opinion. We make suggestions for improving the field of fabrication detection in surveys. We advise a higher degree of skepticism should be applied to surveys from conflict and post-conflict zones than has typically been the case in the past. More work is needed in develop appropriate quality control procedures for situations when conventional methods for detecting interviewer-level data fabrication are insufficient.

Introduction

In September of 2007 the U.S. Broadcasting Board of Governors (BBG) sponsored an opinion poll of 2,000 Iraqis. There were twenty-two supervisors involved in the field work. Of these, six provided data from 684 interviews that display remarkable irregularities. For example, all respondents report that they own a shortwave radio. Yet, on a follow-up question to the 472 eligible respondents who answered "yes" to "Could you please tell me if you have listened to the radio in the last 12 months?" all 472 report that they have never listened to a short wave radio. Answers given to the other sixteen supervisors are much more plausible, ranging over all possible categories for both owning and the number of hours that owners of short wave radios listen to their short wave radios. But for the six identified supervisors all respondents own - but no respondent has ever listened to - a short wave radio.

For this BBG poll short wave radios turn out to provide only one among dozens of anomalies that revolve around these same six supervisors. Moreover, there is a 2008 BBG Iraq poll in which precisely the same six supervisors, identified by the same numbers in both datasets, again generate dozens of suspicious patterns. On top of this, several of these supervisor numbers appear again, presiding over a mass of non-credible data in each of three earlier polls.[1] A single company, KA Research Limited (KARL), performed the fieldwork for all of these polls.

In this paper we explore patterns which suggest widespread fabrication of interviews in five polls conducted in Iraq between 2005 and 2008. We identify the supervisor numbers behind the suspicious data for each of these polls. We refer to these identified supervisors as the "focal supervisors" since we focus on them in our analysis. The evidence that suggests fabrication falls broadly into five types. For categorical variables there are whole categories of responses to particular questions that are not selected at all in any of the interviews of the focal supervisors. A special case of this phenomenon is a common pattern of zero responses for "don't know" and "refused" for the focal supervisors while the other supervisors have positive numbers of responses in these categories. There are a number of implausible relationships between answers to pairs or series of questions, e.g., zero correlations between answers to very similar, even identical, questions asked at different points within a single survey for respondents of focal supervisors. The respondents to the focal supervisors also display restricted ranges in their responses to scale questions relative to the ranges displayed by the respondents to the other supervisors. Finally, television viewing patterns across randomly selected households are, implausibly, tightly interwoven across households for the focal supervisors.

There have been a few related incidents in the survey world in recent years. The owner of the polling firm DataUSA pled guilty in 2006 to falsifying polling data that were provided to the George W. Bush presidential campaign (Blumenthal 2006). Spagat (2010) presents evidence of fabrication in the Burnham et al. (2006) survey of deaths in Iraq after the principal investigator for that survey was censured in 2009 by AAPOR for violating its disclosure standards (AAPOR 2009a) . Strategic Vision LLC was also censured by AAPOR in 2009 (AAPOR 2009b) and subsequently accused of fraud on the influential blog of Nate Silver (Blumenthal 2009). Spagat and Dougherty (2010) found serious irregularities in the data of a separate survey of Iraq deaths (ORB 2008) that are large enough to render the data invalid. The Daily KOS accused the polling firm Research

---

[1] One of these earlier polls was also commissioned by the BBG while the two others were commissioned by the University of Maryland's Program on International Policy Attitudes (PIPA).

2000 of providing them with unsound data, sued and eventually reached a legal settlement with Research 2000 (Kos 2010, Blumenthal 2011). Two census managers working in Brooklyn fabricated roughly 10,000 census returns (Buckley 2010). The existence of so many known recent episodes suggest that fabricated, or at least highly unsound, survey data may be common. Another lesson is that fabrication may enter at various levels of a survey operation, including the level of interview supervisors as has appears to have happened with KARL's Iraq polls.

While working on this paper we were able to draw upon some published ideas on fabrication detection in surveys, although this literature focuses primarily on interviewers rather than supervisors. Murphy et al. (2004) suggest studying "rare response combinations" on the theory that fabricators may overestimate the frequency with which rare answers to particular questions will appear in combination with one another We build on this idea, analyzing responses to combinations of questions not only for rare responses but in general. The idea is that it is easier to fabricate plausible answers to single questions than it is to fabricate plausible relationships between answers to separate questions.   Porras and English (2004) also focus on rare response combinations while adding a second idea that has proved useful for us. Fabricated answers might display smaller variances than legitimate ones because fabricators might avoid extreme responses that can be flagged by quality-control inspectors as outliers requiring further investigation. Porras and English (2004) suggest further that fabricators might underuse the categories of "don't know" and "refuse", in an attempt to avoid detection by making their data look squeaky clean. The five polls that we analyze do display this pattern even though fabricated data examined by Porras and English (2004) do not do so consistently. AAPOR and ASA (2003) lists "off-site isolation of interviewers from the parent organization" as a risk factor for interviewer falsification and, indeed, this seems to be a factor in the five polls we analyze in this paper. In fact, there is a rather complicated chain of responsibility running, in three of the polls, from the U.S. Broadcasting Board of Governors to the InterMedia Survey Institute to US-based D3 Systems to Istanbul-based KARL to KARL's field operation in Iraq.

Our analysis suggests some rethinking of the field of fabrication detection is called for. First, scrutiny should extend beyond just interviewers onto other levels including supervisors. In fact, many of the standard remedies for suspected fabrication, such as interview monitoring or reinterviews, are normally implemented by supervisors. Obviously, these approaches are problematic if supervisors are part of the problem. Second, the current emphasis on outlier detection (e.g., Murphy et al. 2004) misses the possibility that the absence of outliers may also be an indicator of fabrication. In our work we have found the absence of outliers to be a more revealing warning sign than their presence.[2] For example, the most salient feature of the interview durations data for the focal supervisors is that durations are uniformly confined to narrow ranges, looking too perfect to be real.[3] This pattern would be invisible to a fabrication detection method that relies on the presence of outliers.

---

[2] Some of our difference with the standard approach might depend on exactly how outliers are defined. For example, if there are six offered answers to a categorical question and all respondents give the same answer then we could say that there is an outlier. However, this would be a fairly nonstandard usage of the term "outlier" which is generally used only when there is high variance in a distribution rather than zero variance as is the case in this example.

[3] The quality control procedures of D3 Systems (Peng and Field, 2011 & Hornbeck et. al 2010), KARL's US affiliate, may have contributed to the lack of variance and outliers amongst the focal supervisors. For example, these procedures flag interviewers whose average interviewing times deviate by more than one standard deviation from overall averages. Hornbeck et al. (2011) summarize the goal of their procedures as "constant reduction in variability." The focal supervisors appear to be extremely effective according to this criterion.

Third, the issue of fabrication in conflict-zone surveys requires special consideration. Hornbeck et al. (2010), writing for D3 Systems, KARL's US affiliate, stress this point at the outset:

> "Conflict zones present other unique difficulties. It is often impossible for a company to audit or observe the accuracy of face-to-face data collection in these environments by a subcontractor's employees. Yet it is just such environments in which many of the traditional sources of interviewer error, and a few new ones, are most likely to occur. The pressures prompting them can be extreme, up to and including death….In high conflict environments, these interviewers may be tempted to fill out the surveys themselves, and turn them in as completed interviews, in lieu of risking their personal safety." (Hornbeck et al. 2010, p. 1)

Later Hornbeck et al. (2010) suggest that conflict-zone fieldwork should be observed if it is safe enough to do so. From a fabrication-detection perspective there is a problem in that the detection regime is weakest precisely when the risk of fabrication is greatest. Indeed, as we show in the present paper, the areas of Iraq where evidence of fabrication is most prevalent were very dangerous areas during the period covered by the five polls we analyze.

The plan of the paper is as follows. In the next section we discuss the basics of the polling data, including the list of problematic "focal" supervisors and their level of participation in the different polls. The following section presents the evidence of fabrication. Then we outline our conclusions in the final section.

The Data

We obtained data from six public opinion polls conducted in Iraq by KA Research Limited, working with its US affiliate D3 Systems, between March 2005 and November 2008. We used a Freedom of Information Act request to obtain datasets from four polls that were commissioned by the U.S. Broadcasting Board of Governors (BBG). We obtained the other two from the web site of the University of Maryland's Program on International Policy Attitudes (PIPA) which was the client for these polls. We denote each poll by the acronym for the sponsor followed by the month and year of the poll, e.g., BBG-DEC05.[4]

The datasets all include identification codes for interview supervisors (among many other variables). We initially explored interviewer level fabrication as the primary cause of the unusual patterns in the data. However, heavy turnover among interviewers during the period in question caused us to reject this hypothesis. Just 3 of the 45 interviewer codes associated with the problematic data remained consistent across all 5 datasets. Region was also rejected as an explanation, as there were two datasets (BBG-DEC05 and PIPA-JAN06) where half of the Baghdad data does not display the same suspicious patterns[5]. We also considered keypuncher level fabrication, but found no consistency in the keypuncher codes associated with the problematic data. Supervisor code, on the other hand, appeared to be the stable component across datasets, with a consistent, slowly growing set of supervisors accounting for the apparent fabrication present in the datasets. We assume, in reaching this conclusion, that each interviewer, keypuncher, and supervisor code represents a single real person. Supervisor is the highest level individual identified in the dataset, so we also assume there is no higher level person or

---

[4] PIPA-JAN06 is at http://hdl.handle.net/1903/10158 and PIPA-SEP06 is at http://hdl.handle.net/1903/10163. We will post the BBG polls at an appropriate time on a web site.

[5] The half of the Baghdad data which did not show suspicious patterns was from supervisor 42, one of the non-focal supervisors, providing further evidence that the apparent fabrication was at the supervisor level rather than the province level.

people above the supervisor level who are wholly and exclusively responsible for the interviews which show the suspicious patterns.

Following this assessment, we did an initial analysis of each dataset, supervisor by supervisor. The main things we looked for were supervisors whose respondents display restricted ranges for scale variables, partial distributions, i.e., empty categories, for categorical variables, lack of don't know/refused responses, and other highly improbable patterns such as interviews only starting on minutes that are multiples of 5. This exploration led us to identify supervisors 36, 38, 43, 44, 47, 93 and 94 as producing suspicious patterns suggestive of fabrication. These are the ones we refer to as the "focal supervisors". These supervisors turn out to have further suspicious patterns which we also cover below. In particular, we provide evidence that suggests the interviews of the focal supervisors are fabricated in all of the polls in which they participated. Based on both our initial and later, fuller analysis, we do not believe that any of the other supervisors have fabricated entire interviews although, of course, we cannot demonstrate this definitively.

Table 1 displays the participation patterns for the focal supervisors, including the provinces where they operated and the number of interviews they supervised.[6] No focal supervisors were present in BBG-MAR05 and our initial analysis suggests that this poll is fine so we dropped it from our subsequent analysis. The focal supervisors worked in the conflict hot spots of Baghdad, Anbar and Diyala, accounting for all interviews in Diyala and Anbar when they worked there, roughly half of the Baghdad interviews in BBG-DEC05 and PIPA-JAN06 and all the Baghdad interviews in the subsequent three polls. These three governorates account for roughly one third of the population of Iraq and about 70% of the violent deaths recorded in the Iraq Body Count (continuously updated) database during the period when the polls were conducted so this region is a very important part of Iraq. The percent of interviews (unweighted) covered by the focal supervisors ranges from 16% to 44% for the polls they participated in. Thus, the contamination of the data for this central region of Iraq would cause major problems for both local and national estimates.

Table 1. Participation of the Focal Supervisors in the Six Surveys

Supervisor Number:	BBG-MAR05	BBG-DEC05	PIPA-JAN06	PIPA-SEP06	BBG-SEP07	BBG-NOV08
36	NA	Diyala	Diyala	NA	Diyala	Diyala
38	NA	NA	NA	Diyala	NA	NA
43	NA	Baghdad	Baghdad	Baghdad	Baghdad	Baghdad
44	NA	Anbar	Anbar	Anbar	Anbar	Anbar
47	NA	NA	NA	Baghdad	Baghdad	Baghdad
93	NA	NA	NA	NA	Baghdad	Baghdad
94	NA	NA	NA	NA	Baghdad	Baghdad

---

[6] Supervisor 47 also worked on PIPA-JAN06, but in Najaf rather than Baghdad where he/she worked for all subsequent projects. These 33 responses showed evidence of the same partial distribution phenomenon displayed by the focal supervisors in this project, although there are some deviations including the parts of the scale on which these 33 responses were distributed compared to the responses of the other focal supervisors. Because of these deviations and because supervisor 47 was outside of the 3 provinces where the vast majority of the problems occurred, we view this set of 33 responses as an outlier. While they are potentially fabricated we do not include them among the "focal supervisor" cases.

Number of Other Supervisors	NA	17	15	15	16	11
Focal Supervisor Case Count	NA	481	332	442	684	687
Other Supervisor Case Count	NA	2,519	818	708	1,316	876
Total Sample Size	NA	3,000	1,150	1,150	2,000	1,563

Evidence of Fabrication in Five Polls

We divide the evidence of fabrication into five sub-sections: partial distributions of categorical values, empty categories for "don't know" and "refused", improbable correlations between answers to different questions, restricted ranges on scale variables and implausible relationships across households in television viewing patterns.
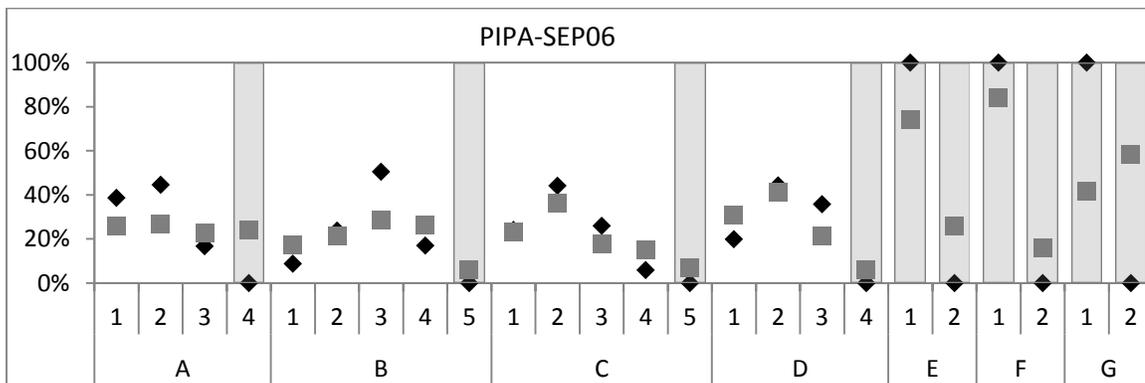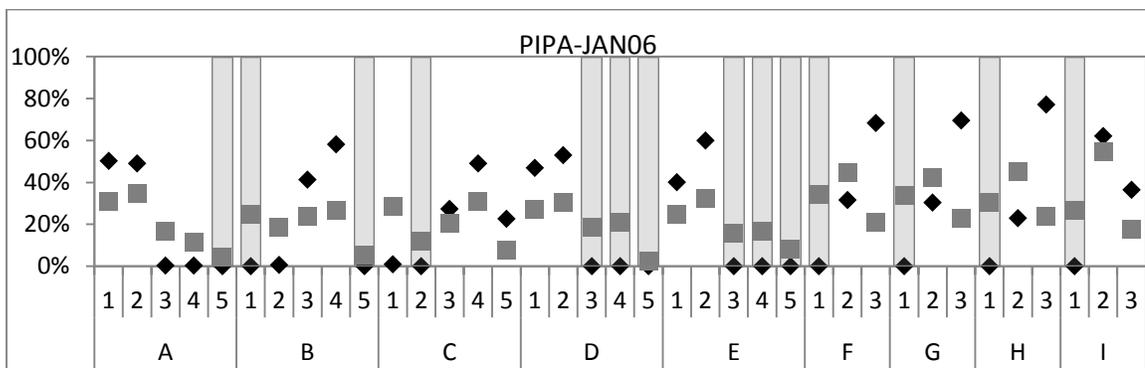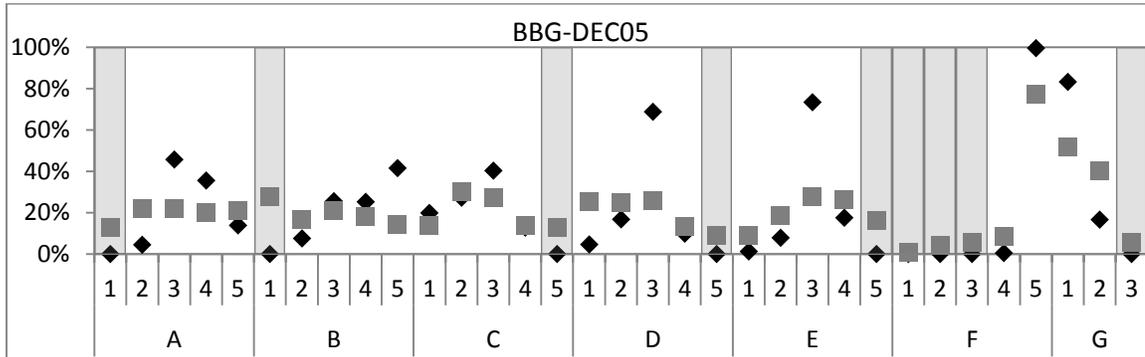
PARTIAL DISTRIBUTIONS OF CATEGORICAL VARIABLES

We assembled 460 categorical variables from the five datasets. These are all the categorical variables we could find after excluding data on the administration of the surveys, data with the potential for a strong regional influence, data constructed from combining or recoding other variables, and variables on the timing of interviewees' media consumption. Excluded regional data cover categories such as language spoken at home, ethnicity, religious affiliation, and use of specific media outlets. Such items are potentially homogeneous within regions, which could confound analysis of supervisor-level variation since supervisors tend to operate within single provinces. Some media outlets in Iraq may serve only single regions or language groups, a factor which may also limit the possible variety of responses.

Figure 1 gives data on a sample of questions (capital letters in the figure) for which responses (numbers in the figure) obtained by the focal supervisors (black diamonds) do not cover the full range of possible responses while the answers obtained by the other supervisors (grey squares) do have full support.[7] The grey shading highlights cases for which either all respondents or no respondents of the focal supervisors answer in a certain way while the respondents' answers for the other supervisors are well away from 0% or 100%. For example, for BBG-DEC05 0 out of 481 respondents for the focal supervisors is very favorably inclined toward Turkey (question A) compared to 328 out of 2,519 respondents for the other supervisors (see appendix for full question text). The BBG-SEP07 poll provides three extreme cases (questions C, D and E) for which there are six possible choices yet all 684 respondents for the focal supervisors choose only one response; they never use foreign satellite TV, international radio or text messaging for information about current events. More than a third of the respondents for the other supervisors choose different responses. Responses to such questions are

---

[7] We give the list of questions and allowable answers in the appendix for all the questions covered in figure 1.

subject to numerous sources of error[8], making the notion of a uniform response across 684 respondents thoroughly implausible.

BBG-DEC05

PIPA-JAN06

PIPA-SEP06

[8] Roger Tourangeau, Lance J. Rips, Kenneth A. Rasinski, "The psychology of survey response"
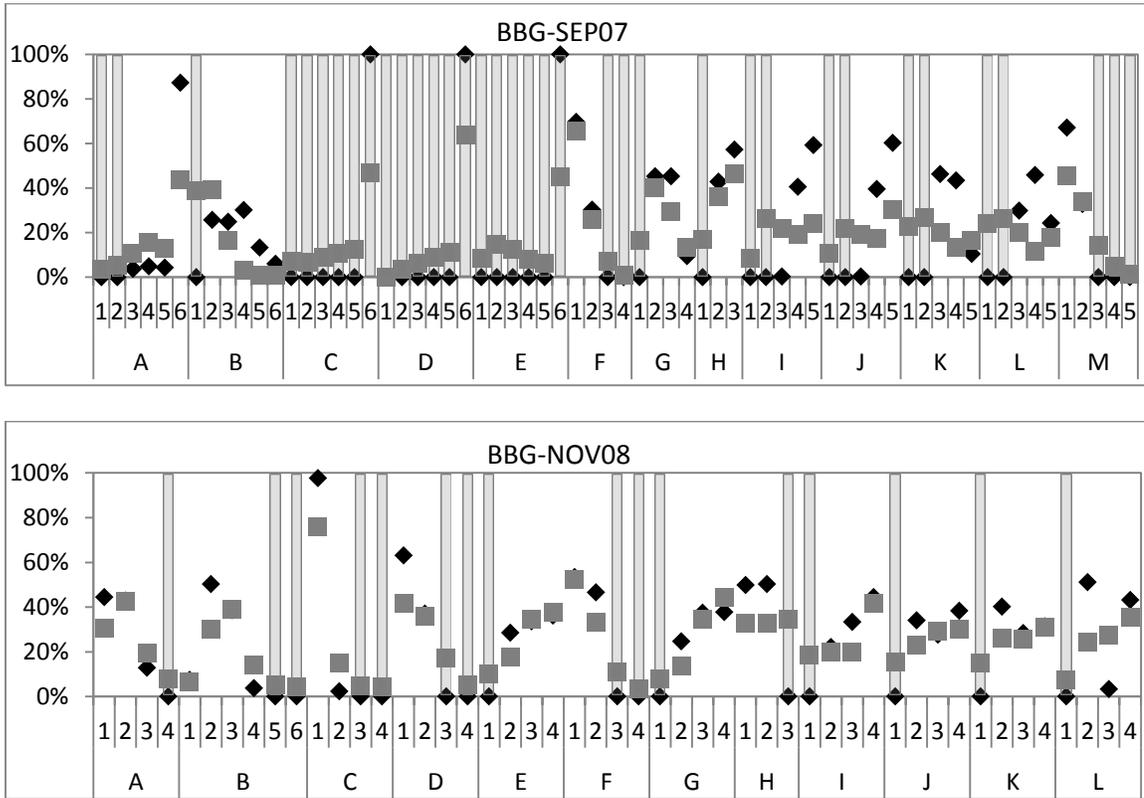
Figure 1. Focal Supervisors Report Partial Distribution for Many Categorical Variables

Note: Questions are represented by capital letters, categories of responses by numbers, the focal supervisors are the black diamonds and the other supervisors are the grey squares. The grey-shaded strips highlight questions for which either 0 or all respondents to the focal supervisors are recorded as selecting a particular category while some but not all of the respondents for the other supervisors select that category. It sometimes looks to the naked eye like there should be shading when in fact the responses are slightly inside the interval [0,100].

It is rare in opinion polling for whole categories to be unanimously passed over by a large number of respondents. Complete unanimity on one category out of six offered categories is almost unheard of. Even when there is a high unanimity of opinion amongst potential respondents there is still, normally, at least a little variation in answers. The mental processes respondents undertake when answering survey questions can be viewed as having multiple steps, each of which can contribute to variability in responses, something which is frequently lacking for the focal supervisors. Tourangeau, Rips and Rasinski (2000) present a useful, four step cognitive model of cognition, retrieval, judgment, and response. Each step of the process contributes variation to survey response, undermining the empty categories displayed in figure 1. Misinterpretation of questions, imperfect recall, erroneous inference, differing interpretations of the scales offered for responses are among the many factors that add variation to question responses, even among individuals who disagree about very little [Tourangeau. Rips and Rasinski (2000)]. There will be still greater variation to the extent that people actually disagree with one another - as they often do. Further error, and hence variation, enters real data through such factors such as limited interviewer comprehension, poor transcription and mistakes in data entry [Biemer and Lyberg (2003)]. The empty response categories that are frequently present among the focal supervisors, and conspicuously absent among the rest, are inconsistent with such known realities of survey research.

For a few highly politicized questions the population in the hot spots of Baghdad, Anbar and Diyala might exhibit a uniformity of responses that is unusual in opinion polling. However, many of the questions with either completely empty and/or completely full categories for the focal supervisors are entirely non-political. For example, all 442 respondents for the focal supervisors own a shortwave radio, a radio, and a cell phone (PIPA-SEP06); not one of the 481 respondents thinks friends will influence whom they will marry or where they will work (BBG-DEC05); and none of the 684 respondents reports being "very interested" in news about the U.S., Europe, other Arab countries, or "any other country" (BBG-SEP07).

On some questions, zero respondents for the focal supervisors tick a neutral response while positive numbers give weak opinions on both sides of neutral. For example, on PIPA-SEP06 among the 442 focal-supervisor respondents 105 (24%) think inter-ethnic violence will "increase a little" while another 75 (17%) think it will "decrease a little" but not a single one thinks that US withdrawal will have "no effect either way." Of the 708 respondents to the other supervisors the numbers are, respectively, 152 (21%), 187 (26%) and 42 (6%).

We can take the distributions for the other supervisors as "true" distributions and use these to calculate probabilities that the respondents to the focal supervisors will display the distributions they do, that is assuming the responses for the focal supervisors are drawn from the distributions for the other supervisors. These probabilities almost always turn out to be virtually zero. For example, 13% of other supervisors' respondents to question A for BBG-DEC05 give response 1 which 0 out of 481 respondents give for the focal supervisors. The calculated probability for the focal supervisors to have an empty category in this case is 0.87 raised to the power 481, which is zero to many decimal places. In fact, even if we arbitrarily lower the "true" probability of response 1 by an order of magnitude to 0.013 the probability of all 481 respondents for the focal supervisors choosing 1 is still only 0.002. So we can easily reject a hypothesis that the unanimity for the focal supervisors arises due to pure chance. And recall that question A on BBG-DEC05 is just one such question among many with this suspicious pattern. Any reasonable compound probability of so many empty categories will be zero for any practical purpose.

Table 2 summarizes the information on all of the 460 categorical variables we studied. We employ the following definitions. There is a partial distribution on a particular question if at least one scale point contains a zero response.[9] Table 2 also shows that the respondents of the other supervisors never display partial distributions at the aggregate level while the focal supervisors, on the other hand, display partial distributions a total of 171 times across the five datasets.

---

[9] We exclude the responses "Don't Know", "Refused", "Not Sure", and "Ineligible" from this accounting. Including these would raise the counts of partial distributions for focal supervisors much higher without having a big impact on the counts for the other supervisors.

Table 2. Partial Distributions for Focal Supervisors and the Other Supervisors

		Dataset (Count of categorical variables)				
		BBG DEC05 (167)	PIPA JAN06 (57)	PIPA SEP06 (60)	BBG SEP07 (93)	BBG NOV08 (83)
Partial distribution present at aggregate level	Focal Supervisors	54	23	14	49	31
	All Others	0	0	0	0	0

"EMPTY CATEGORIES OF "DON'T KNOW" AND "REFUSED""

Fabricators may try to avoid detection by making their data look particularly clean, although they may err on the side of making them look too clean. A manifestation of this temptation could be to avoid answers of "don't know" or "refused". Indeed, Table 3 shows that the focal supervisors frequently had zeros for both of these categories on questions for which the other supervisors enter positive numbers for both. The reverse never happens. The most extreme case is PIPA-JAN06 for which on 41 questions out of 57, the focal supervisors have zero "don't know" and "refused". The focal supervisors give zeros for both "don't know" and "refused" for the majority of questions for which we know that these categories were offered. [10]

---

[10] The last row of Table 3 gives the number of questions that we are not able to resolve unambiguously given the data we have. In these cases no supervisor, focal or otherwise, records any "don't know" or "refused" but it is not clear whether these categories were actually offered to respondents during survey administration.

Table 3. Use of "Don't Know" and "Refused" Categories

		Dataset Count of Questions				
		BBG DEC05	PIPA JAN06	PIPA SEP06	BBG SEP0	BBG NOV08
		167	57	60	93	83
Exclusive non-use of Don't Know / Refused categories	Problem Supervisors	50	41	37	42	45
	All Others	0	0	0	0	0
Both groups used Don't Know / Refused Categories		47	5	4	2	9
Subtotal: Don't Know / Refused categories appears at least once		*97*	*46*	*41*	*44*	*54*
Don't Know / Refused Categories not used		70	11	19	49	29

IMPROBABLE CORRELATIONS BETWEEN ANSWERS TO DIFFERENT QUESTIONS

A common technique for fabrication detection is to look at sets of questions for which certain combinations of responses are rare [Murphy et al. (2003) & Porras and English (2004)]. The idea is that it is difficult for fabricators to understand all the complex interrelationships that might exist between separate questions.   In this section we generalize this principle, testing for predictable correlations between question responses beyond just rare combinations.

Table 4 gives a pair of questions from each poll that should yield positively correlated answers. The correlations in the table are for Arab Sunni respondents only, to isolate the correlations to examine only within-group variation. For BBG-DEC05 and BBG-SEP07 the two questions are nearly identical to one another but appear far apart in the sequence of questions, perhaps as a test of question wording. In the other cases, the questions are a little different but one would still expect many people to answer them similarly. In line with these expectations, the responses to interviews conducted under the umbrella of the other supervisors are indeed strongly correlated. On the other hand, the responses for the focal supervisors are only weakly correlated, contrary to expectations. This pattern is consistent with fabrication of responses question by questions, i.e., in isolation from one another, rather than coordinating the responses across questions. Of course, sophisticated fabricators would produce positive correlations between these pairs of questions if they focused on them specifically. However, it is hard to anticipate all likely relationships between questions to cover all likely correlations.

Table 4. Pairs of Questions that Fail to Display Expected Correlations for Focal Supervisors

Survey	Questions	Correlations	
		Focal supervisors	Other supervisors
BBG-DEC05	Early in questionnaire: "How important is it to you, personally, to stay informed about news and current events?"	-0.04	0.37
	Late in questionnaire: "How interested are you, personally, to stay informed about current events in Iraq?"		
PIPA-JAN06	"If US-led forces withdraw from Iraq in the next six months, do you think presence of foreign fighters in Iraq will increase a lot, increase a little, decrease a little, decrease a lot, or have no effect either way?"	.05	0.63
	"If US-led forces withdraw from Iraq in the next six months, do you think the amount of crime will increase a lot, increase a little, decrease a little, decrease a lot, or have no effect either way?"		
PIPA-SEP06	"Nouri Maliki's government doing a good job in its efforts to deal with Iraq's problems?"	-0.04	0.47
	"Please tell me if you have a very favorable, somewhat favorable, somewhat unfavorable or very favorable opinion of:...Prime Minister Nouri Al-Maliki?"		
BBG-SEP07	Early in questionnaire: "How interested are you in staying informed about current events?"	0.01	0.42
	Late in questionnaire: "Generally speaking, how interested would you say you are in current events?"		
BBG-NOV08	"How favorably or unfavorably inclined are you personally towards USA?"	0.08	0.57
	"How favorably or unfavorably inclined are you personally towards Great Britain?"		

Sometimes correlations are too strong. For example, all 131 respondents associated with Supervisor 43 in PIPA-JAN06 answered the following questions identically, with 60 responding "increase a lot" both times and 71 responding "increase a little" both times.

- "If US-led forces withdraw from Iraq in the next six months, do you think **day-to-day security for ordinary Iraqis** will increase a lot, increase a little, decrease a little, decrease a lot, or have no effect either way?"
- "If US-led forces withdraw from Iraq in the next six months, do you think **availability of public services such as electricity, schools, sanitation, etc.** will increase a lot, increase a little, decrease a little, decrease a lot, or have no effect either way?"

An examination of longer sequences of questions provides further suspicious patters which suggest fabrication. Sometimes there are sequences of questions, each with the same answers never ticked for the focal supervisors. For example, in BBG-NOV08 the questionnaire includes a battery of favorability ratings on a four point scale toward the UK, U.S., Germany, France, Russia, and Iran. For the last 5 nations in this sequence, no respondent of the focal supervisors chooses the "very favorable" option.
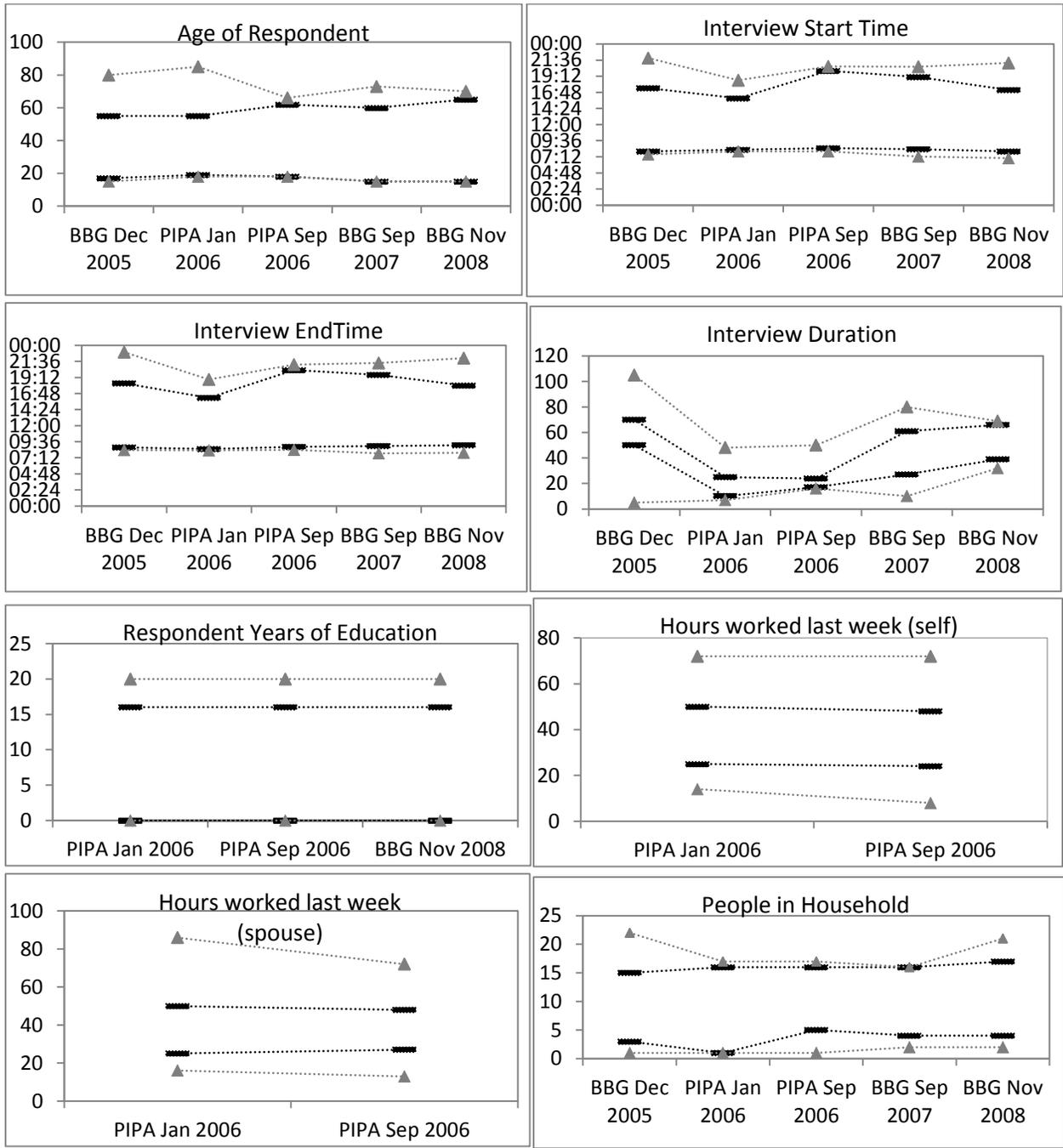
A similar dynamic occurs in BBG-SEP07. In this instance, favorability for 4 countries is measured on a 5 point scale, and none of the 684 respondents uses either of the top two scale points. This is not a reflection of uniformly unfavorable views of the countries in question, as there are a considerable number of middle-category responses of "neither favorable nor unfavorable" for France (316 respondents or 46%) and for Germany (204 respondents or 30%). It is unrealistic that there would be so many neutral responses without a single person rising to the next category of "rather favorably inclined".

RESTRICTED RANGES ON SCALE VARIABLES

Porras and English (2004) argue that fabricating interviewers "would attempt to escape detection by avoiding outliers". They provide evidence that this actually happens, using some data known with certainty to be fabricated. On the other hand, outlier detection is a commonly used strategy in surveys to uncover fabrication in surveys [Murphy et. al (2003), AAPOR and ASA (2003)]. Outliers are tricky to handle for fabricators because each one can draw attention and suspicion, yet real data do contain outliers; so the absence of outliers can also be suspicious, albeit less obviously than their presence. Porras and English (2004) focus their detection efforts on variances of responses. However, we believe that the presence or absence of outliers can be identified more directly through an analysis of ranges. We provide such an analysis in this section.

Figure 3 shows the ranges for all scale variables appearing in the five polls. The black dashes are for the focal supervisors and the grey triangles designate the other supervisors. The ranges for the focal supervisors are contained within the ranges for the other supervisors in every case. Sometimes the former are well within the

latter. This relative lack of variation for the focal supervisors is suggestive of fabrication with the fabricators lacking the ability or imagination to predict outliers or full scale ranges. [11]

Age of Respondent

Interview Start Time

Interview EndTime

Interview Duration

Respondent Years of Education

Hours worked last week (self)

Hours worked last week (spouse)

People in Household

---

[11] Since the other supervisors handled more interviews than the focal ones we might expect some tendency for narrower ranges for the latter compared to the former. However, relatively small numbers for the focal supervisors certainly cannot explain extremely different ranges for variables such as hours worked or domestic workers in the household.
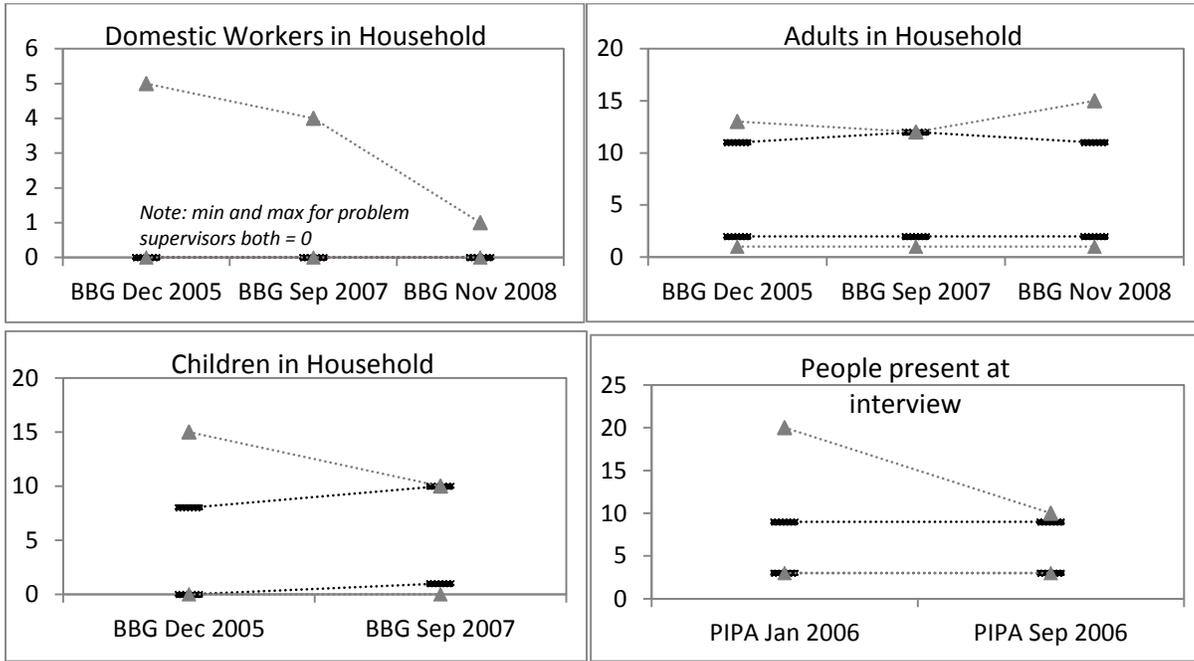
Figure 3. Responses for Focal Supervisors Display Restricted Ranges Compared to the Other Supervisors

Note: Focal supervisors are black dashes and all other supervisors are grey triangles.

In figure 4 we explore full distributions, rather than just ranges as we do in figure 3. The figure displays the distributions over interview durations in all five polls, both for the focal supervisors and the other supervisors. The distributions for the focal supervisors have truncated tails, which are generally more prominent on the right. For the BBG-DEC05 poll, fully 98% (after rounding) of the interviews are coded as lasting 50, 55 or 60 minutes for the focal supervisors. For PIPA-JAN06 all but 1 interview fall between 10 and 17 minutes, while 53% of interviews for other supervisors fall outside this range. For PIPA-SEP06 every single one of the interviews of the focal supervisors last between 17 and 24 minutes whereas for the other supervisors 46% of the interviews are outside this narrow range. BBG-NOV08 looks better than the others, although it still has a truncated left tail. Yet another strange fact is that all the interviews of the focal supervisors in the BBG-DEC05 dataset start at times that are exact multiples of 5. This percentage declines over the course of the subsequent four polls (99%, 69%, 40% and 34%) perhaps suggesting increasing sophistication at fabricating distributions that appear plausible.
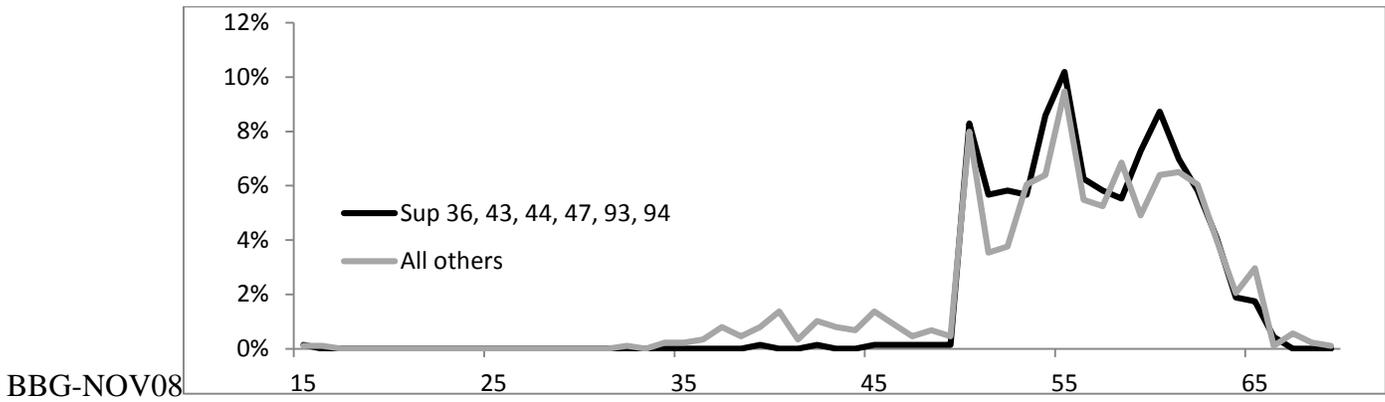
BBG-DEC05

PIPA-JAN06

PIPA-SEP06

BBG-SEP07

17
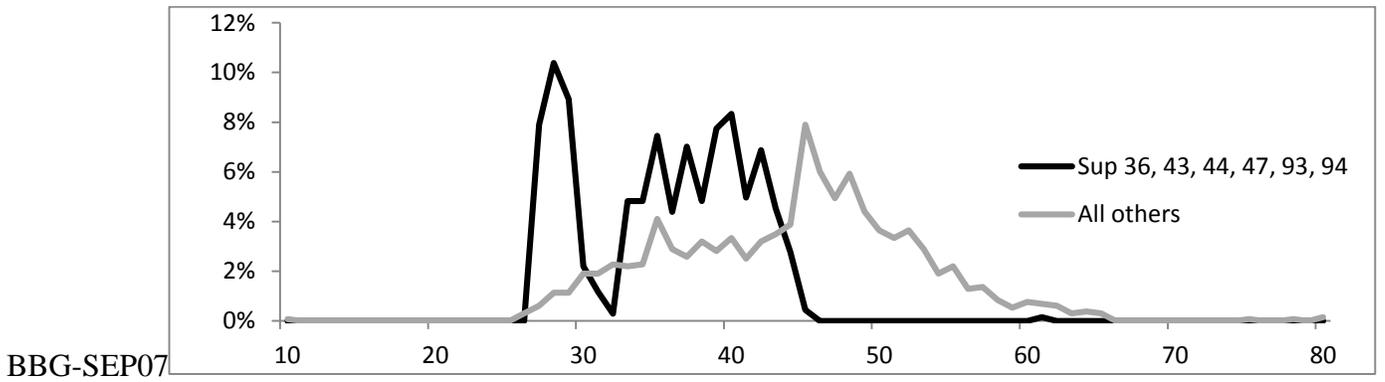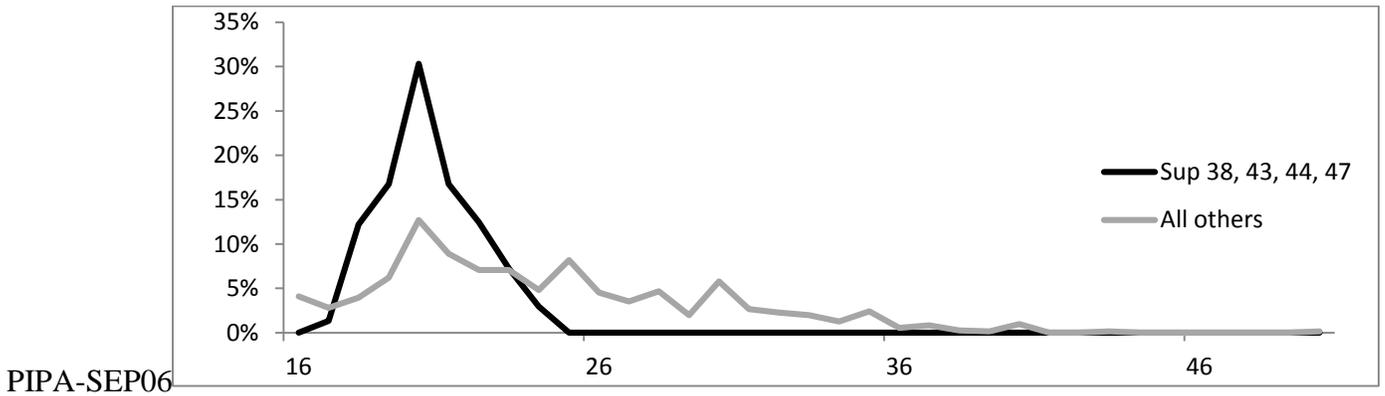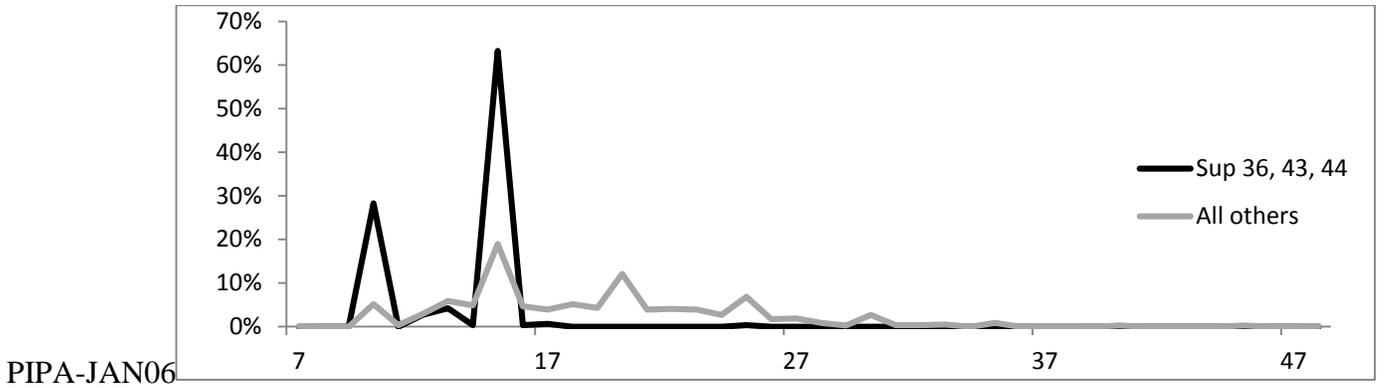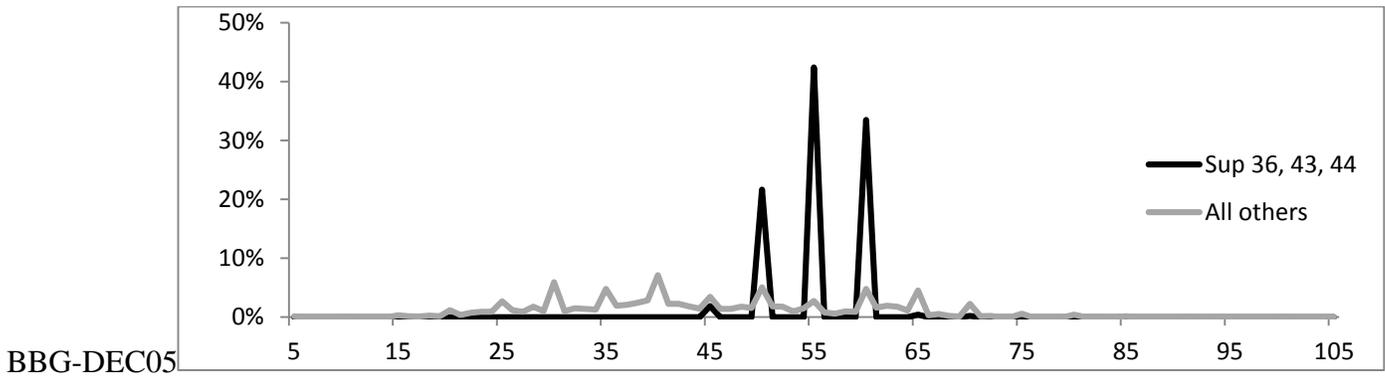
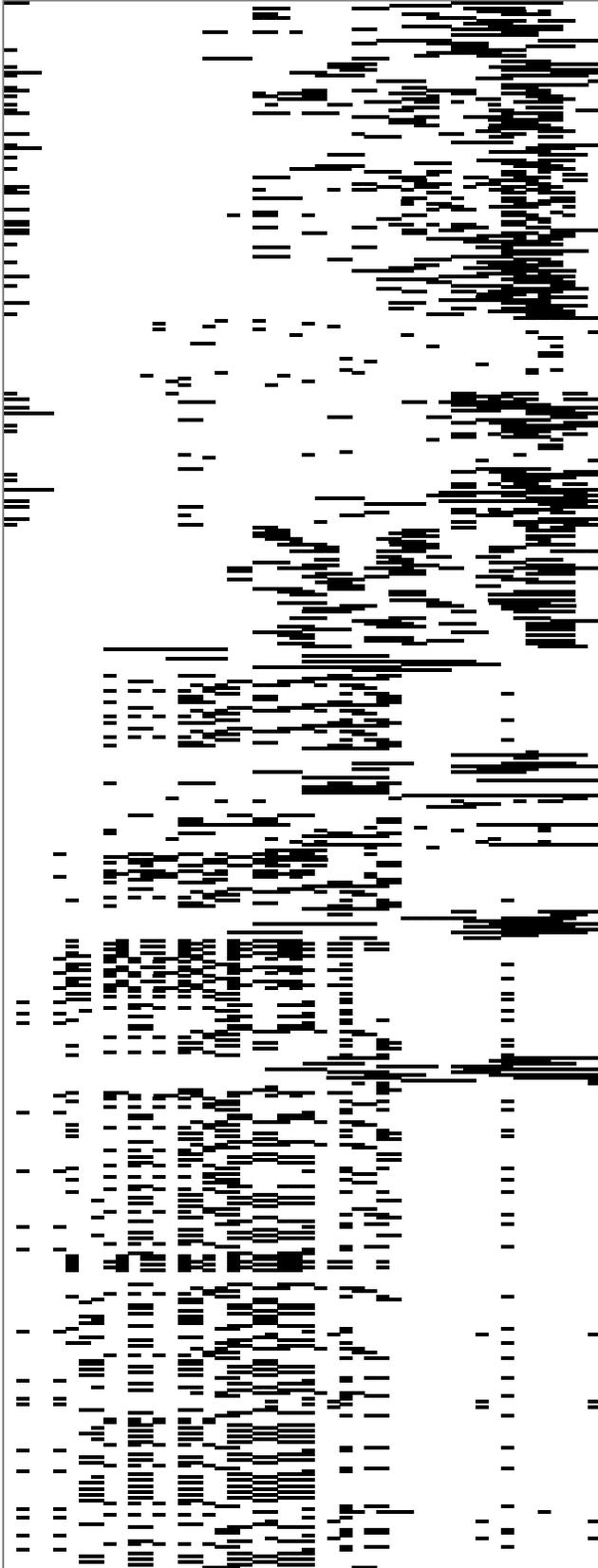**Figure 4. The Distributions of Interview Durations**

IMPLAUSIBLE RELATIONSHIPS ACROSS HOUSEHOLDS IN TELEVISION VIEWING PATTERNS

Figure 5 displays television viewing patterns recorded in the BBG-DEC05 poll for all of the focal supervisors on the right and for a random sample of the other supervisors on the left.[12] Each row represents the viewing pattern for a particular interviewee for the previous day. The day is divided into 48 half-hour blocks with blocks filled for the time slots for which the respondent says that he/she watched television. The rows are ordered by the supervisor variable and then the "serial number" variable, i.e. the unique identifier given to each record in the BBG data files.   The first line is serial number 1200 of supervisor 36 followed by serial 1201 of supervisor 36. The cases for supervisor 43 begin immediately after the last case for supervisor 36. These are, again, presented in ascending order of serial numbers. We finish with supervisor 44. The interviews of the other supervisors in the left-side block are ordered similarly.

We would expect households to spend big blocks of time with their televisions on and other blocks with it turned off. Indeed, we see much of this pattern for the other supervisors (on the left). The respondents for the focal supervisors frequently switch their televisions on for half-hour slots before shutting them off. In some areas, mainly in the top half of the right-hand picture, these half-hour blocks seem quite uniform across households. Even more striking are the diagonal streaks in many areas of the right block. These patterns suggest implausibly tight relationships between the different respondents for the focal supervisors.

---

[12] Data from BBG-SEP07and BBG-NOV08 look similar.

Random Selection of Other Supervisors   Focal Supervisors
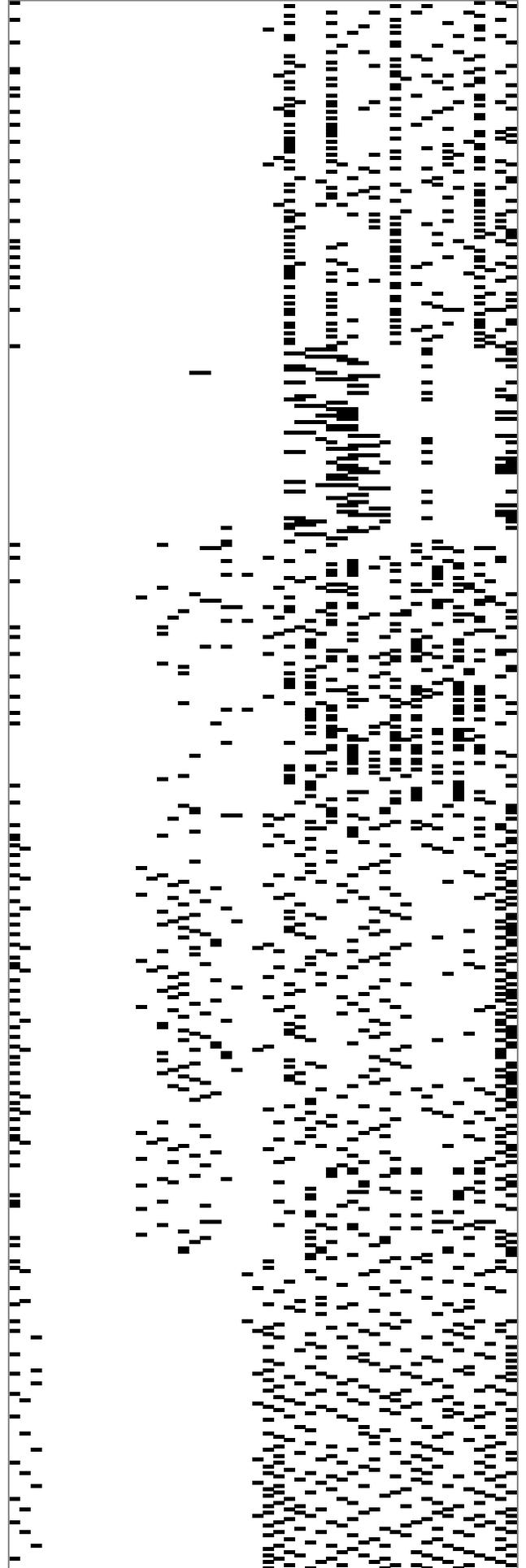
Figure 5. Television Viewing Patterns for Other Supervisors (left) and Focal Supervisors (right)

Figure 6 contains an extract from figure 5 drawn from BBG-DEC05, supervisor 44. There is a diagonal streak in which the first household switches on the television at 20:30 and switches off at 21:00. The next household of the same supervisor then switches on at 21:00 and off at 21:30. The next household for that supervisor switches on at 21:30 and off at 22:00. The pattern continues down a string of seven households for this supervisor. Such patterns recur again and again in figure 5. Yet these households are supposed to be randomly selected and they are certainly not selected based on their relationships with one another regarding their television viewing patterns. We can think of no plausible mechanism that could bring about such tightly knit relationships.[13]
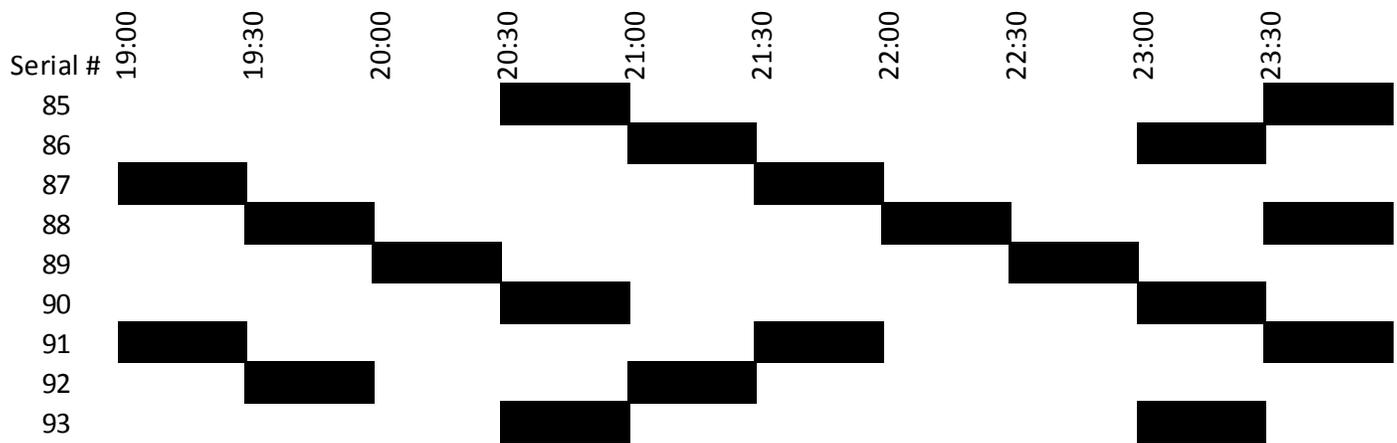

Figure 6. An Extract from Figure 5

Conclusion

The data associated with the focal supervisors in the five polls shows many suspicious patterns which suggest the strong possibility of fabrication. Given the high percentage of cases covered by the focal supervisors, between 16% and 44% depending on the poll, if our suspicious are confirmed, any national figures derived from these datasets must also be considered unsound. The figures for Baghdad, Anbar and Diyala appear completely unusable for these five polls.

KA Research Limited has collected Iraq data on behalf of numerous organizations besides PIPA and the BBG. These clients include ABC News, the BBC, the Japan Broadcasting Corporation (NHK), the U.S.

---

[13] We note that some of the patterns for the other supervisors also appear to be strange, particularly at the bottom of figure 5 where there seem to be some recurring patterns. The question on television viewing times is complicated and some other supervisors may have fabricated their responses in order to save some time.

Department of State (FedSpending.org 2009), and the International Research Exchange (IREX and D3 Systems 2010). It is important that these and other consumers of KARL public opinion data, particularly Iraq data, do their own assessments to see whether their data also shows evidence of potential fabrication.

KARL's extensive and high-profile client list ensures that KARL-produced data has played a significant role in shaping international understanding of Iraqi public opinion. However, if the patterns we have found in five KARL polls extend beyond PIPA and the BBG then much of the collected body of knowledge of Iraqi public opinion must now be reconsidered.

It appears that U.S. foreign policy may have been informed by a considerable quantity of flawed data. PIPA's Iraq data have been reported widely, and briefed to Congress on at least one occasion (Kull 2008). While the specifics are not reported regarding how BBG's data was used within the U.S. Government, it stands to reason that it had some impact on U.S. actions during the period in question. If fabrication extends beyond these cases, particularly to polls commissioned by the State Department then the damage would be still greater.

The apparent fabrication becomes obvious in the five Iraq polls only when we view the data through the proper lens – focal supervisors versus the other supervisors. Without this lens the patterns remain murky. Moreover, the literature on fabrication in surveys does not point in the right direction in this case because this literature focuses on interviewers rather than supervisors. In fact, many fabrication detection and correction mechanisms, such as reinterviewing randomly selected or suspicious cases, would probably have been ineffective in KARL's case as these procedures are normally implemented by supervisors who are at the heart of the problem in the five polls we have analyzed. Although the datasets contain two variables which indicate interviews that were subject to "back check", the system described by these variables could also have been a fabrication, or else was totally ineffective at detecting and rooting out the apparent problem. Thus, we do not think that the fabrication in the five Iraq polls was easy to detect, despite its great scale.

Additionally, the literature on fabrication detection tends to assume a higher level of technical sophistication during data collection than is typically available in post-conflict environments. Interviews are often conducted by paper and pencil in these environments, rendering impossible analysis of time between questions, GPS tracking of interviewers, and other methods used in domestic face to face survey research.

One obvious lesson from our analysis is that fabrication detection systems should account for the possibility of fabrication at all levels in a survey, including supervisors, data entry people, interviewers, and anyone else in the chain of reporting of poll results. Related to this is the possibility that multiple layers of sub-contracting might be a risk factor for fabrication. PIPA, for example, hired D3 Systems which worked with KARL which is based in Istanbul but has field presence in Iraq. With so many layers it is hard to ensure accountability.

Fabrication detection in surveys is roughly analogous to steroid detection in sports; it is a two-way competition in which it is hard for the detectors to stay ahead of the cheaters. Good detection will always involve both art and science since new cases are likely to differ from old ones.

It is, of course, extremely difficult to conduct a high-quality survey in a conflict zone. It is probably no accident that the suspicious patterns showed up in the most violent region of Iraq. Interviewers would have to risk their lives to work in this zone. We suggest that surveys conducted in war zones should provoke healthy

skepticism and should have unusually strong fabrication-detection and quality-control procedures in place. While such quality control measures will likely increase costs and time to field surveys, this case shows that extensive quality control is a cost that cannot safely be avoided. If anything, quality control has typically been lower in conflict zones than in peaceful places, suggesting that the apparent fabrication we have uncovered in this paper may only be the tip of an iceberg.

References

AAPOR. 2009a. "AAPOR Finds Gilbert Burnham in Violation of Ethics Code, http://www.aapor.org/Content/aapor/Resources/ForMedia/PressReleases/AAPORFindsGilbertBurnhaminViolationofEthicsCode/default.htm.

AAPOR 2009b. "AAPOR Raises Objections to Actions by Atlanta-Based Strategic Vision LLC." http://www.aapor.org/Content/aapor/Resources/ForMedia/PressReleases/AAPORRaisesObjectionstoActionsbyAtlantaBasedStrategicVisionLLC/default.htm.

AAPOR and ASA. 2003. "Interviewer Fabrication in Survey Research," http://aapor.org/AM/Template.cfm?Section=Report_to_AAPOR_Standards_Comm_on_Interviewer_Falsification1&Template=/CM/ContentDisplay.cfm&ContentID=3160Biemer, Paul P. and Lyberg, Lars E. 2003. Introduction to Survey Quality Hoboken NJ, Wiley and Sons.

Blumenthal, Mark. 2006. "The Pollster That Made Up Data." Pollster.com. http://www.pollster.com/blogs/the_pollster_that_made_up_data.php?nr=1.

Blumenthal, Mark. 2009. ""Strategic Vision Ends Year in Deep, Deep Hole," National Journal, December 21, http://www.nationaljournal.com/njonline/mp_20091221_3652.php.

Blumenthal, Mark. 2011. "Daily KOS versus Research 2000 Lawsuit Settled." HUFFPOST POLLSTER http://www.huffingtonpost.com/2011/05/27/daily-kos-research-2000-lawsuit_n_867775.html.

Broadcasting Board of Governors (BBG-MAR05)). 2005a. Data File Obtained through Freedom of Information Act, Washington DC.

Broadcasting Board of Governors (BBG-DEC05), 2005b.   Data File Obtained through Freedom of Information Act, Washington DC.

Broadcasting Board of Governors (BBG-SEP07), 2007. Data File Obtained through Freedom of Information Act, Washington DC.

Broadcasting Board of Governors (BBG-NOV08), 2008. Data File Obtained through Freedom of Information Act, Washington DC.

Buckley, Cara. 2011. "Two Brooklyn Census Managers Fired Over Faked Surveys." The New York Times, July 8, http://cityroom.blogs.nytimes.com/2010/06/26/two-brooklyn-census-managers-fired-over-faked-surveys/.

Burnham, G. Lafta, R. Doocy, S. and Roberts, L. 2006. "Mortality after the 2003 invasion of Iraq: a cross-sectional cluster sample survey, The Lancet 368 (9545) 1421-1428.

FedSpending.org. 2009. "Contracts to Contractor(s) 'ka research limited' (FY 2006)," OMB Watch, http://www.fedspending.org/fpds/fpds.php?fiscal_year=2006&company_name=ka+research+limited&sortp=u&datype=T&reptype=l&database=fpds&detail=4&submit=GO.

Hornbeck, Stephen, Peng, David, Studner, Catherine and Jodice, David. 2010. "Ensuring Data Quality in Conflict Zones." D^3 Systems, http://www.d3systems.com/wp-content/uploads/2010/09/ensuring_data_quality_in_conflict_zones.pdf.

Iraq Body Count. Continuously updated. Available at http://www.iraqbodycount.org/.

IREX and D3 Systems. 2010. "Iraq Media Study – National Audience Analysis." IREX, http://www.irex.org/resource/iraq-media-study-national-audience-analysis.

Kos. 2010. "Research 2000: Problems in Plain Sight," Daily Kos, http://www.dailykos.com/story/2010/6/29/880179/-Research-2000:-Problems-in-plain-sight.

Kull, Steven. 2008. "Iraqi Public Opinion on the Presence of US Troops." World Public Opinion.org, http://www.worldpublicopinion.org/pipa/articles/brmiddleeastnafricara/517.php?nid=&id=&pnt=517.

Murphy, Joe, Baxter, Rodney, Eyerman, Joe, Cunningham, David and Kennet, Joel. 2004. "A System for Detecting Interviewer Falsification." RTI International, http://www.rti.org/pubs/paper-detectingintrvfalsifi.pdf.

ORB. 2008. "Update on Iraqi Casualty Data," http://www.opinion.co.uk/Newsroom_details.aspx?NewsId=120.

Peng, David and Field, Karl. 2011. "Quality Control in Telephone Survey Interviewer Monitoring, Survey Practice, April: www.surveypractice.org.

Porras, Javier and English, Ned. 2004. "Data-Driven Approaches to Identifying Interviewer Data Falsification: The Case of Health Survyes. In JSM Proceedings, Survey Research Methods Section. Toronto, Canada: American Statistical Association. 42232-4228.

Program on International Policy Attitudes (PIPA-JAN06). 2006a. "Iraq Poll Illuminates Sunni-Shia Divide: Sunni Dissatisfaction Widespread," http://drum.lib.umd.edu/handle/1903/10158.

Program on International Policy Attitudes (PIPA-SEP06). 2006b. "Most Iraqis Want U.S. Troops Out Within a Year: Say U.S. Presence Provoking More Conflict Than it is Preventing Approval of Attacks on U.S.-led Forces Rises to 6 in 10." http://drum.lib.umd.edu/handle/1903/10163.

Spagat, Michael. 2010. "Ethical and Data-Integrity Problems in the Second Lancet Survey of Mortality in Iraq." Defense and Peace Economics 21(1) 1-41.

Tourangeau, Roger, Rips, Lance J. and Rasinski, Kenneth. 2000. The Psychology of Survey Response, Cambridge university Press, Cambridge, U.K.

Spagat, Michael and Dougherty, Josh. 2010. "Conflict Deaths in Iraq: A Methodological Critique of the ORB Survey Estimate." *Survey Research Methods* 4(1) 3-15.

Appendix

This appendix contains all the questions, together with their answers, that are covered in figure 1 of the main paper. We use italics below when zero respondents for the focal supervisors gave these answers. We use bold below when all respondents for the focal supervisors gave these answers.

BBG-DEC05[14]

A. How favorably or unfavorably inclined are you personally towards Turkey?
1. *Very favorably inclined*
2. Somewhat favorably inclined
3. Neither favorably nor unfavorably inclined
4. Somewhat unfavorably inclined
5. Very unfavorably inclined
B. How favorably or unfavorably inclined are you personally towards Great Britain?
1. *Very favorably inclined*
2. Somewhat favorably inclined
3. Neither favorably nor unfavorably inclined
4. Somewhat unfavorably inclined
5. Very unfavorably inclined
C. How favorably or unfavorably inclined are you personally towards Saudi Arabia?
1. Very favorably inclined
2. Somewhat favorably inclined
3. Neither favorably nor unfavorably inclined
4. Somewhat unfavorably inclined
5. *Very unfavorably inclined*
D. How favorably or unfavorably inclined are you personally towards France?
1. Very favorably inclined
2. Somewhat favorably inclined
3. Neither favorably nor unfavorably inclined
4. Somewhat unfavorably inclined
5. *Very unfavorably inclined*
E. How favorably or unfavorably inclined are you personally towards Russia?
1. Very favorably inclined
2. Somewhat favorably inclined
3. Neither favorably nor unfavorably inclined

---

[14] Actual question text not available for BBG-sponsored surveys. Question text shown here is drawn from the SPSS files provided in response to the FOIA request.

4. Somewhat unfavorably inclined
5. *Very unfavorably inclined*
F. How often use community meetings to get information about current events?
   1. *Daily or almost every day*
   2. *Several times a week*
   3. *Once a week*
   4. Less than once a week
   5. Never
G. Thinking about a News and Information TV Channel that would serve the whole family, how important would it be for this TV Channel to offer the following? Day long coverage of a major news story.
   1. Very important
   2. Somewhat important
   3. *Not at all important*

PIPA-JAN06

A. If US-led forces withdraw from Iraq in the next six months, do you think willingness of factions in parliament to cooperate will increase a lot, increase a little, decrease a little, decrease a lot, or have no effect either way?
   1. Increase a lot
   2. Increase a little
   3. Decrease a lot
   4. Decrease a little
   5. *No effect either way*
B. If US-led forces withdraw from Iraq in the next six months, do you think Amount inter-ethnic violence will increase a lot, increase a little, decrease a little, decrease a lot, or have no effect either way?
   1. *Increase a lot*
   2. Increase a little
   3. Decrease a lot
   4. Decrease a little
   5. *No effect either way*
C. If US-led forces withdraw from Iraq in the next six months, do you think presence foreign fighters in Iraq will increase a lot, increase a little, decrease a little, decrease a lot, or have no effect either way?
   1. Increase a lot
   2. *Increase a little*
   3. Decrease a lot
   4. Decrease a little
   5. No effect either way
D. If US-led forces withdraw from Iraq in the next six months, do you think day to day security for ordinary Iraqis will increase a lot, increase a little, decrease a little, decrease a lot, or have no effect either way?
   1. Increase a lot
   2. Increase a little
   3. *Decrease a lot*
   4. *Decrease a little*
   5. *No effect either way*

E. If US-led forces withdraw from Iraq in the next six months, do you think Availability of public services will increase a lot, increase a little, decrease a little, decrease a lot, or have no effect either way?
   1. Increase a lot
   2. Increase a little
   3. *Decrease a lot*
   4. *Decrease a little*
   5. *No effect either way*
F. I am going to read a list of how the US has been involved in Iraq in non-military ways. For each one that I read, please tell me if you approve of the US being involved and think the US is doing a good job, approve but think the US is doing a poor job, or disapprove of the US being involved in Iraq in this way....assisting with economic development of Iraq?
   1. *Approve and US is doing a good job*
   2. Approve, but US is doing a poor job
   3. Disapprove
G. I am going to read a list of how the US has been involved in Iraq in non-military ways. For each one that I read, please tell me if you approve of the US being involved and think the US is doing a good job, approve but think the US is doing a poor job, or disapprove of the US being involved in Iraq in this way....assisting with the development of Iraq's oil industry?
   1. *Approve and US is doing a good job*
   2. Approve, but US is doing a poor job
   3. Disapprove
H. I am going to read a list of how the US has been involved in Iraq in non-military ways. For each one that I read, please tell me if you approve of the US being involved and think the US is doing a good job, approve but think the US is doing a poor job, or disapprove of the US being involved in Iraq in this way....helping Iraqis organize their communities to address local needs?
   1. *Approve and US is doing a good job*
   2. Approve, but US is doing a poor job
   3. Disapprove
I. Would you prefer US or UN take the lead in Iraq's economic reconstruction?
   1. *US*
   2. UN
   3. Neither

PIPA-SEP06

A. Do you strongly approve, approve somewhat, disapprove somewhat, or strongly disapprove of the following:...Attacks on US-led forces in Iraq?
   1. Strongly approve
   2. Approve somewhat
   3. Disapprove somewhat
   4. *Strongly disapprove*
B. If US-led forces withdraw from Iraq in the next six months, do you think the amount of inter-ethnic violence will increase a lot, increase a little, decrease a little, decrease a lot, or have no effect either way?
   1. Increase a lot
   2. Increase a little

      3. Decrease a lot
      4. Decrease a little
      5. *No effect either way*

C. If US-led forces withdraw from Iraq in the next six months, do you think the day to day security for ordinary Iraqis will increase a lot, increase a little, decrease a little, decrease a lot, or have no effect either way?
      1. Increase a lot
      2. Increase a little
      3. Decrease a lot
      4. Decrease a little
      5. *No effect either way*

D. How likely do you think Iraq will be a single state from 5 years from now?
      1. Very Likely
      2. Somewhat Likely
      3. Not Very Likely
      4. *Not At All Likely*

E. Household Own...Cell Phone?[15]
      1. Yes
      2. *No*

F. Household Own...Radio?
      1. Yes
      2. *No*

G. Household Own... Shortwave radio?
      1. Yes
      2. *No*

BBG-SEP07

A. How often use Magazines to get information about current events?
      1. *Daily or almost every day*
      2. *Several times a week*
      3. 1-2 days a week
      4. From time to time in the last month
      5. Less often
      6. Never use

B. How often use Friends or family members to get information about current events?
      1. *Daily or almost every day*
      2. Several times a week
      3. 1-2 days a week
      4. From time to time in the last month
      5. Less often
      6. Never use

C. How often use Foreign Satellite TV Channels to get information about current events?
      1. *Daily or almost every day*

---

[15] For this questionnaire, the cell phone, radio, and shortwave ownership variables were included in the demographics. The questions for these variables were not included on the questionnaire found online at http://drum.lib.umd.edu/handle/1903/10163

  2. *Several times a week*
  3. *1-2 days a week*
  4. *From time to time in the last month*
  5. *Less often*
  6. Never use

D. How often use International Radio to get information about current events?
  1. *Daily or almost every day*
  2. *Several times a week*
  3. *1-2 days a week*
  4. *From time to time in the last month*
  5. *Less often*
  6. Never use

E. How often use SMS (Text Messaging via mobile phone) to get information about current events
  1. *Daily or almost every day*
  2. *Several times a week*
  3. *1-2 days a week*
  4. *From time to time in the last month*
  5. *Less often*
  6. Never use

F. In general, how interested are you in News about your country?
  1. Very interested
  2. Fairly interested
  3. *Not very interested*
  4. *Not at all interested*

G. In general, how interested are you in News about Arab countries?
  1. *Very interested*
  2. Fairly interested
  3. Not very interested
  4. Not at all interested

H. How interested you are in watching Business?
  1. *Very interested*
  2. Somewhat interested
  3. Not at all interested

I. How favorably or unfavorably inclined are you personally towards Great Britain?
  1. *Very favorably inclined*
  2. *Rather favorably inclined*
  3. Neither favorable nor unfavorable
  4. Rather unfavorably inclined
  5. Very unfavorably inclined

J. How favorably or unfavorably inclined are you personally towards USA?
  1. *Very favorably inclined*
  2. *Rather favorably inclined*
  3. Neither favorable nor unfavorable
  4. Rather unfavorably inclined
  5. Very unfavorably inclined

K. How favorably or unfavorably inclined are you personally towards France?
  1. *Very favorably inclined*
  2. *Rather favorably inclined*
  3. Neither favorable nor unfavorable

4. Rather unfavorably inclined
5. Very unfavorably inclined
L. How favorably or unfavorably inclined are you personally towards Germany?
   1. *Very favorably inclined*
   2. *Rather favorably inclined*
   3. Neither favorable nor unfavorable
   4. Rather unfavorably inclined
   5. Very unfavorably inclined
M. Generally speaking, how interested would you say you are in current events?
   1. Very interested
   2. Interested
   3. *Somewhat interested*
   4. *Not so interested*
   5. *Not at all interested*

BBG-NOV08

A. How interested are you, personally, in staying informed about current events in Iraq?
   1. Very interested
   2. Somewhat interested
   3. Not very interested
   4. *Not at all interested*
B. In general, how often do you access any type of news and current affairs? By any, I mean news and current affairs on radio, TV, internet, newspapers etc...?
   1. Every hour
   2. 2-3 times a day
   3. Once a day
   4. 2-3 times a week
   5. *Once a week*
   6. *Less often*
C. How often do you use television (any) to get information about current events?
   1. Every day of the week
   2. At least once a week
   3. *Less often*
   4. *Never*
D. How often do you use Arab satellite TV to get information about current events?
   1. Every day of the week
   2. At least once a week
   3. *Less often*
   4. *Never*
E. How often do you use International TV to get information about current events?
   1. *Every day of the week*
   2. At least once a week
   3. Less often
   4. Never
F. In general how interested are you in: News about your country?

1. Very interested
2. Somewhat interested
3. *Not very interested*
4. *Not at all interested*

G. In general how interested are you in: News about Europe?
   1. *Very interested*
   2. Somewhat interested
   3. Not very interested
   4. Not at all interested

H. Are you interested in seeing a TV program on Documentaries on current affairs?
   1. Very interested
   2. Somewhat interested
   3. *Not at all interested*

I. How favorably or unfavorably inclined are you personally towards USA?
   1. *Very favorably inclined*
   2. Somewhat favorably inclined
   3. Somewhat unfavorably inclined
   4. Very unfavorably inclined

J. How favorably or unfavorably inclined are you personally towards Germany?
   1. *Very favorably inclined*
   2. Somewhat favorably inclined
   3. Somewhat unfavorably inclined
   4. Very unfavorably inclined

K. How favorably or unfavorably inclined are you personally towards France?
   1. *Very favorably inclined*
   2. Somewhat favorably inclined
   3. Somewhat unfavorably inclined
   4. Very unfavorably inclined

L. As you look towards the future, which system do you think is best suited to Iraq?
   1. *The restoration of the Ba'athist government*
   2. An Islamic government based on Shari'a law
   3. A decentralized, federal and secular democratic republic
   4. A centralized and secular democratic republic