

Listening Survey Summary

There were 9 panelists and each was allowed to listen to the various streams in any order they chose. All 9 panelists reviewed the Iraq stream and the Levant stream. Due to time constraints, only 5 panelists reviewed the Morocco stream.

GRAMMAR AND SYNTAX

There were significant problems with grammar, syntax and pronunciation from many of the broadcasters on all 3 streams. The names of President Bush, Vice-President Cheney, and the Financial Times were frequently mispronounced. Even some Arabic names and Arabic words were mispronounced. Many panelists agreed that the poor pronunciation would likely be detected by an Arab and result in questions about the broadcaster's ability. One panelist commented that there seem to be some broadcasters who are making most of the mistakes because they have very little experience and lack good training. A few panelists felt that the syntax and pronunciation were acceptable, but only 2 of the 9 panelists did not indicate any problems with broadcasters using appropriate syntax and pronunciation.

Occasionally, the English was translated incorrectly and the wrong Arabic word was used. For example, "overpayment" got translated into a complex set of words, which made it unclear and difficult to understand the message. Another example occurred with the word supplies, which is "muarid" in Arabic, but was translated to "mawrid" instead. "Mawrid," which means resources in Arabic, was wrongly translated on the script and then read by several broadcasters. Others examples noted were the Arabic word meaning "to review or correct" instead of the word for "to amend a law" in the Levant stream, and the Arabic word for "to shift" was used instead of the word for "to move" in the Iraq stream. A serious mistranslation occurred with a sound byte of President Bush, in which he said "spend well," but the Arabic used meant "behave rationally."

Five of the panelists commented about a problem with the use of wrong prepositions. The use of wrong prepositions is likely due to the use of dialects instead of modern standard Arabic by the broadcasters. In addition, some serious grammatical errors such as improper plural form, lack of gender agreement, and lack of adjective-noun agreement were noted; however, most of the panelists felt that in general the grammar was acceptable.

CLARITY AND STYLE

Most of the panelists rated the overall clarity and appropriateness of the stories and newscasts, along with style of presentation as "Rather Clear." A higher rating was not assessed due to issues with the broadcaster's style and technical problems. The Iraq stream had a mean score of 3.61 (sd 0.928). The Levant stream had a mean score of 3.61 (sd 0.928). The Morocco stream had a mean score of 3.38 (sd 1.109).

Comprehension of the presentation was impaired by the broadcaster's delivery. A few of the panelists felt that the broadcaster was understandable but could be improved by reading more slowly. Three of the respondents noted that the broadcasters in all 3 streams frequently read too fast, which made comprehension difficult. One broadcaster uses a very monotone voice, which makes the presentation less interesting. It was further noted that the use of dialects may cause some groups to have difficulty understanding the message clearly. Many of the panelist agreed that the main broadcaster in Iraq was nearly perfect, but the stringers were not very impressive. One panelist added a comment that the presentation does not even come close to other Arabic radio stations.

Most of the panelists felt that the broadcasts were not offensive to any person or group; however, there were some serious incidents noted. The call to prayer heard in the background of the Iraq stream broadcast would be offensive to Muslims. Another potential offensive incident was that Egyptian President Mubarak was mentioned in a broadcast without using his title "President," which is important in Egyptian culture. It was also noted that Egyptian and Lebanese dialects dominated the broadcast on the Iraq stream. Further, the absence of a Moroccan dialect is noticeably absent among the presence of other dialects on the Moroccan stream. A panelist explained that "it is condescending to use Levant terminology on the Moroccan stream." Some Arabs may also be offended by the names of celebrities being mispronounced on the Levant stream.

The transmission quality on all 3 streams was a mixture of good and poor. There is a hissing sound and echoing on the broadcast streams, which seems to occur mostly when there's background music being played. The Moroccan stream seems to have the loudest hissing sound and most echoing, but it was also occasionally noticed on the Iraq stream and the Levant stream. In addition, the background music on the Iraq stream occasionally overpowered the voice of the broadcaster, which resulted in the drowning out of words. A very irritating sound of the broadcaster breathing could also often be heard on the Iraq stream and the Levant stream.

GENERAL IMPRESSION AND OBSERVATIONS

Most of the panelists rated the stories, special programming, and newscasts as good overall. The Iraq stream had a mean score of 4.50 (sd 0.907). The Levant stream had a mean score of 4.39 (sd 0.993). The Morocco stream had a mean score of 4.00 (sd 1.000). A rating for the Morocco stream was only provided by 3 panelists.

The panelists did not respond to the question about special programs correctly, so an accurate statistical analysis could not be accomplished. It can be inferred from the responses that there were likely no call-in shows on the Levant or Morocco streams; none of the panelists indicated any call-in shows and a few noted that there were absolutely none. Only 1 panelist indicated a call-in show on the Iraq stream; however, some noted that there were none. Similarly, only 2 panelists noted an editorial on the Levant stream, and only 1 panelist indicated that there was an editorial on the Iraq and Morocco streams, with a few indicating that there were none. The majority of panelist indicated that there were interviews and news magazines or public announcements, but some noted that there were none.

The capture of former Iraqi President Saddam Hussein was discussed for the entire news period on the first hour following his arrest and then frequently broadcasted on all 3 streams, but the delivery was average overall. Most of the panelists indicated that the poor grammar and syntax used by the broadcasters degraded the effectiveness of the message. One panelist noted that the Baghdad reporter covering the story was not very good. In addition, there was some static and background noise detected in the broadcasts.

Most of the panelists declared that there was no "breaking news" because all of the news broadcasts were provided at the scheduled time. A few panelists indicated that there were one or two "breaking news" broadcasts; however, these panelists further explained that was provided during the scheduled time, but seemed to be treated as significant because interviews were included in those broadcasts. Due to the disparity of what would be defined as "breaking news" no accurate statistical analysis could be done. It can be inferred from the responses, however, that the news was done at scheduled times and no programming was interrupted to broadcast "breaking news" such as the capture of former Iraqi President Saddam Hussein.

Most of the panelists did not seem to understand how to answer this question, so no accurate statistical analysis could be conducted with the data. The responses indicate, however, that there were some incidents of repeated news broadcasts in the same day. One panelist noted that it was mostly in the late night parts of the broadcast day.

ADDITIONAL COMMENTS

A few panelists commented that Radio Sawa does not seem to be anything unique and special that would draw an audience away from other news radio broadcasts in the Middle East. Further, they explained that Radio Sawa does not seem to have a clear-cut mission. One panelist also commented that "Radio Sawa fails to present America to its audience."

Some of the panelists also provided recommendations for improvement. First, reading the Arabic news script more slowly will likely improve the problems with pronunciation of Arabic words; however, more training, is needed for broadcasters to improve their pronunciation of names and English words. Second, increasing the variety of news stories and programming would help to improve Radio Sawa's attractiveness. Finally, more local news for each broadcast area should be included and dialects should reflect the broadcast area if they are going to be used on the radio.

**APPENDIX B: Possible Factors Contributing to the Differences
in the Pew and ACNielsen Survey Results**

OIG contracted a statistical and survey research expert to examine the collection and validity of audience research data used to evaluate the credibility and impact of Radio Sawa programming, including data compiled by ACNielsen and Pew. Survey results for these two research organizations noted differences in respondents' favorability toward the United States. OIG's contractor was asked to examine the possible factors contributing to differences between the ACNielsen and Pew survey data.

The Pew Global Attitudes Project is a worldwide public opinion survey conceived "to gauge attitudes in every region...and to measure attitudes towards democracy and other key issues." After the terrorist attacks on September 11, 2001, the project was broadened to measure attitudes about terrorism, the intersection between Islamic faith and public policy in countries with significant Muslim populations, and to more deeply probe attitudes towards the United States in all countries.

ACNielsen's study, conducted to create a baseline for future assessment of Radio Sawa's impact, was based on interviews conducted during the first two weeks of April 2002 among 1,150 randomly selected adults over the age of 15 in Amman, Jordan. The survey asked questions about media and music habits and preferences, trust in media, interest in news and current affairs, Internet access and use, and attitudes toward the West. The OIG contractor provided the table below and the discussion that follow on possible factors contributing to differences between the two survey results.

**Draft
UNCLASSIFIED**

In the Pew study, nine nations were surveyed; Jordan was the only nation that was also in the ACNielsen study. Table D below is OIG's contractor comparison of the results regarding respondents' favorability toward the United States.

Table D: Comparison of the Survey Results in Jordan Regarding Respondents' Favorability Toward the United States

	ACNielsen	Pew	
Favorability	July-August 2003 ^a	May 2003	March 2004
Favorable	21.6%	1%	5%
Neither favorable nor unfavorable	15.4%	Not applicable	Not applicable
Somewhat unfavorable	8.1%	16%	26%
Very unfavorable	54.9%	83%	67%

^a Excluding refusals and don't know comments.

Source: OIG contractor analysis from BBG data and Pew data.

OIG's contractor found several factors that might contribute to the differences in the survey results. First, the differences in sampling strategy could produce differences. If any of the samples was skewed toward a certain subpopulation group, the survey estimates could be systematically biased with regard to favorability. For example, if the ACNielsen sample contained a higher proportion of younger people (based on weighted data) than the Pew sample, the survey estimates from the ACNielsen study could be more positive than those from the Pew study. Because the Pew report only mentioned that the survey was based on "a nationally representative sample," there is inadequate information to make comparisons between the sampling strategies of the two studies and to identify their differences.

Second, the order of the questions in the survey questionnaires could be another contributing factor. In the Pew surveys, the favorability question was the second question asked, soliciting peoples' spontaneous responses. In the ACNielsen survey, however, the favorability question was placed toward the end of the questionnaire,

Draft
UNCLASSIFIED

following all of the questions about respondents' foreign media use (e.g., Radio Sawa and BBC) and their opinions about these media. These preceding questions could serve to build up some pro-Western sentiment and lead to more favorable opinions about the United States.

Third, the wording of the favorability question was not identical in the two surveys. In the Pew survey, the question was worded: "Please tell me if you have a very favorable, somewhat favorable, somewhat unfavorable, or very unfavorable opinion of the United States." In the ACNielsen survey, the question was worded: "Some people like certain countries more than others. Thinking specifically about the United States, how favorably or unfavorably inclined are you personally toward the U.S?"

Again, the Pew question was more likely to solicit spontaneous responses whereas the ACNielsen question prompted more thinking by respondents before answering it. Moreover, by using the words "inclined" and "personally," the ACNielsen question is less direct than the one used in the Pew study.

Fourth, the response categories were different for the favorability questions used in the two studies. In the ACNielsen study, the response categories are "very favorably inclined," "somewhat favorably inclined," "neither favorable nor unfavorable," "somewhat unfavorably inclined," and "very unfavorably inclined." The Pew study did not have the neutral category, nor the word "inclined." In the ACNielsen study, 15.4 percent of respondents chose the neutral category. Most may have held an unfavorable feeling toward the United States but did not want to express it. Without the neutral category, they might have been forced to choose one of the unfavorable categories.

Finally, the difference in the timing of the surveys might also contribute to the differences in the opinions because any change in the environment could cause change in people's opinions. In this case, the differences between the Pew and Nielsen studies would reflect a true change.

**Draft
UNCLASSIFIED**

APPENDIX E Evaluation of the Radio Sawa Research Survey

The analyses of OIG's contracted statistical and survey research expert and an OIG statistician³³ identified numerous inaccuracies and discrepancies in the Radio Sawa Research Survey design and results. The key findings of the two experts, from BBG's presentation, *Alhurra and Radio Sawa: Progress Report, April 2004*, are presented below.

Sample Bias

Selection Bias. BBG's presentation claims the survey data collected by InterMedia in July-August 2003 and February-March 2004 represents the whole of the six studied countries for Radio Sawa. However, in four countries 100 percent of the sample had been selected from urban areas and in one of the remaining two countries 90 percent of the sample was selected from an urban area. This means the survey results could be biased, due to under representation of rural areas. The BBG presentation does not explain why the same countries were not sampled for both Alhurra (television) and Radio Sawa. Also, it is unclear why the BBG surveyors sampled different countries for Alhurra and Radio Sawa.

Self-selection Bias. Listeners who are more disposed to trust the radio stations they listen to may be inflating Radio Sawa's credibility ratings. However, BBG's surveyors can control for this by comparing Radio Sawa's ratings with those of the BBC, Radio Monte Carlo, and one or more of the most listened to local stations in each country, among listeners with similar background characteristics, such as age, gender, and education. In addition, the listeners used for comparison should be those who listened to the relevant

³³ The contractor had 16 years of experience in program evaluation and survey methodology while working on more than 50 federal, state, commercial, and international projects and the OIG statistician had 27 years of experience as an auditor and statistician.

**Draft
UNCLASSIFIED**

radio stations in the past 12 months, including both current and past listeners. The inclusion of multi-station listeners who tune in to Radio Sawa, the BBC, and the most popular local station would also control for the effect of listeners' predisposition to trust the station they listened to.

Imprecise Data Use

Unweighted Averages. The use of unweighted averages in the BBG survey appears inappropriate because of the great difference in population size between the nations studied. Using a weighted mean may be more accurate. In addition, the BBG surveyors should change the unweighted average for "non-listeners" titled "Views of the USA among Radio Sawa Listeners and Non-Listeners" from its incorrect calculation of 24 percent to 22 percent.

Mixed Sampling Results. According to BBG's April 2004 presentation, the surveyors used 2004 underlying Radio Sawa data for five of the countries and 2003 underlying data for the remaining country (Qatar), creating a time discrepancy of over six months. Mixing sampling results in this manner can undermine the quality of data. Also, closer scrutiny of sample sizes allotted to the six studied countries shows that four countries' numbers are from 2003, one country's numbers are from 2004 (Morocco), and one country's numbers cannot be matched to data from either year (Kuwait). As a result, this information, found at the last page of the BBG presentation, contradicts the assertion on the first page of the presentation that data for the five nations arose from 2004 numbers.

Missing Categorization Data. BBG's presentation (April 2004) graphically depicts information on how reliable Radio Sawa listeners believe news and information is on a weekly basis. However, the underlying data for the presentation does not categorize this information on a weekly basis, but instead on a 12-month basis, and this distinction is important to the survey as a whole. Also, BBG's presentation discussed how favorably Radio Sawa listeners and non-listeners are inclined toward the United States. However,

**Draft
UNCLASSIFIED**

the underlying data apparently does not break out this information between listeners and non-listeners.

Missing Confidence Level. BBG's April 2004 presentation indicates a margin of error of plus or minus 2.9 percent, but this information is meaningless without a stated confidence level.

Causal Inferences Versus Correlation

Unrecognized Reasons for Improvement. High numbers or improvements in Radio Sawa's listenership could be attributed to improved content of the station or to other factors, including improved promotion and improved reception. For example, 95 percent of Moroccan respondents indicated in the BBG survey that Radio Sawa reception is either excellent or good, but only 41 percent responded similarly for BBC London.

Inadequate Causal Link Between Rates of Listenership and Views of the United States. To properly measure Radio Sawa's impact in this area, BBG research needs to employ longitudinal survey data to examine differential rates of change in Radio Sawa listeners' views towards the United States and discover if there is a greater increase in favorability towards the United States from baseline point to the follow-up point among weekly listeners compared to annual listeners, or among annual listeners compared to non-listeners, controlling for baseline characteristics (such as age, gender, education, relevant attitudes and behaviors).

Inadequate Causal Link Between Radio Sawa Broadcasting and Progress in the War Against Terrorism. To appropriately gauge Radio Sawa's impact in this area, BBG research needs to measure repeatedly respondents' attitudes and behaviors towards terrorism and the war against terrorism, controlling for respondents' baseline characteristics and environmental factors.