Remediation Decision-Making and Behavioral Economics: Results of an Industry Survey

by Wilson S. Clayton

Abstract

Decision methods applied in the remediation industry were evaluated using a survey of industry practitioners, in order to assess the relative roles of quantitative decision analysis and gut intuition. Principals from the disciplines of behavioral economics and decision theory were used as a framework to evaluate remediation decision behaviors revealed by the survey. The survey was completed by 118 respondents representing academia, consultants, clients, and others. Survey questions focused on perceptions and experiences related to inputs to decisions and decision processes, as well as remediation goal setting and outcomes. The most common remediation objective cited was short-term interim measures and the least common was no further action (NFA) with clean closure. NFA was also sparingly achieved: 33% of respondents reported zero NFA closures in their career, and an overall 15% to 20% NFA closure rate was reported among more experienced respondents. Data inputs were ranked most important to decisions, while the decision process itself was ranked lowest. Intuition-based decision methods such as asking for a trusted opinion, rules of thumb, and meetings were all used at last twice as often as decision analysis such as discounted cash flow or probabilistic analysis. Analysis of survey responses showed that cognitive biases, including overconfidence effect and intuition bias, are present to some extent in remediation decision-making. Practitioners are advised to be mindful of the decision-making processes they apply, and to incorporate elements of both intuition and decision analysis, as appropriate to the decisions being made.

Introduction

Remediation management of subsurface soil and groundwater contamination involves numerous complex decisions related to setting goals and strategy, selecting and designing technologies, determining appropriate performance monitoring, and more. The remediation industry is a science and engineering based field, and decision-making is typically backed up by data, calculations, and detailed analysis. However, gut intuition based on prior experience, expert advice, rules of thumb, and other heuristics also frequently guide the decision-making process. This paper presents the results of a detailed survey of remediation industry practitioners that was designed to assess remediation decision-making practices, and in particular to assess the relative prevalence of quantitative decision analysis vs. gut intuition-based practices. The survey presented herein also sheds important light on remediation goals established and outcomes attained, as reported by survey participants.

Background

The value of structured decision-making has long been recognized in fields such as medicine (Stiell 2000), aeronautics (FAA 2008), and the military (U.S. Army 2015). Similarly, the regulatory framework around most remediation programs, such as the U.S. Environmental Protection Agency (U.S. EPA) Superfund program, consists of a structured multistep process that lends structure to the decision-making that is required. Within these regulatory frameworks, industry practitioners make decisions on the collection of site characterization data, performance of bench-scale and field-scale pilot testing, and performance of detailed analysis such as risk assessments, calculations, and modeling. From all of this information, we decide upon site-specific remediation goals, timeframes, and other requirements. Ultimately, we decide upon the selected remediation approach by comparing alternatives using multiple criteria such as effectiveness, cost, and stakeholder acceptance. While the U.S. EPA and similar regulatory frameworks establish an overall framework to remediation decisions, these systems do not generally address the specific decision-making tools and methods that are applicable to remediation management.

Before we consider decision-making on remediation projects, it is worthwhile to review the background on
decision-making, in general. Although most decisions are made with a mixed set of decision tools, there are essentially two ends of the decision-making spectrum: gut-intuition and decision analysis, and we will consider them separately.

At the quantitative end of the spectrum, decision analysis includes use of ranking algorithms, quantitative decision tools, and sometimes probabilistic methods to address uncertainty in outcomes (Buchanan and O’Connell 2006). Ben Franklin is attributed with the first use of quantitative decision-making (Franklin, 1772), having introduced a quantitative algorithmic decision process based on the now familiar Pro & Con list. In the modern era, engineering economics, or discounted cash flow analysis using net present value (NPV) calculations (Stermole and Stermole 2009), is a common quantitative decision tool focusing on financial metrics. Multicriteria decision analysis (MCDA) tools, pioneered by Keeney and Raiffa (1976), are able to consider nonfinancial metrics, and formalize the balancing of tradeoffs, risk, and uncertainty when faced with multiple decision criteria. Probabilistic decision methods, such as decision trees (Raiffa 1968), account for uncertainties in the probabilities and costs of various possible outcomes of each decision alternative. Decision trees, often with sensitivity analysis and Monte Carlo simulation, are commonly applied to evaluate investments involving high uncertainty such as oil and gas development (Schuyler and Newendorp 2013).

For environmental remediation, decision-making guidance documents are available that commonly outline criteria for site-specific cleanup standards and remedy requirements (e.g., Oregon DEQ 2003, and others), but which provide limited focus on the use of decision analysis. The U.S. EPA has developed remedy decision support tools (U.S. EPA 2005), which focus on developing decision inputs such as risk assessments, models, statistical analysis, remedy screening, and associated tools. The Triad approach (Crumbling et al. 2001) has been widely adopted in the remediation industry as an improvement in decision-making. The Triad approach focuses on project planning, adaptable work strategies, and real-time measurement in addition to laboratory chemical data collection. In terms of quantitative decision analysis, Kiker et al. (2005) reviewed case studies of the application of MCDA in environmental decision-making and showed that it has powerful utility, especially where multiple stakeholders with different decision criteria are involved. They concluded, however, that “formal applications of MCDA in the management of contaminated sites are rare at present.” The U.S. EPA has been in the process of developing a MCDA tool (U.S. EPA 2011); however it is not currently available. Although not in widespread use, decision trees and Monte Carlo analysis of multiple possible decision outcomes have been applied for site remediation decisions (Favara 2010) including probabilistic estimation of corporate environmental liabilities across multiple sites (Stegman 2014).

At the gut-intuition end of the decision spectrum, the human behaviors around decision-making have been studied and reported extensively, and excellent summaries are provided by the Nobel Prize winning behavioral economist Kahneman (2011) and the popular author Gladwell (2005). These authors and others have shown that human behaviors preferentially tend toward gut intuition-based decision-making as opposed to more quantitative analytical decision-making. These authors also showed that humans commonly use heuristics, or mental shortcuts in decision-making, such as rules of thumb, common sense, educated guesses, relying on past experience, and getting an expert opinion. Herein, we will consider gut intuition to include both knowing something as true without proof or analysis, as well as the use of heuristics. Gladwell (2005) shows how the rapid snap judgment (i.e., gut intuition) of an expert can sometimes be more accurate than detailed analysis. Gut intuition decisions, informed by adequate experience, can be very effective, especially in emergencies or when adequate time and/or information is not available for decision analysis (Buchanan and O’Connell 2006; Fox 2014). Although gut intuition can lead to accurate decisions, it also can lead to poor decisions because it is subject to cognitive biases which form decision traps (Kahneman 2011).

The seminal work on cognitive biases (Tversky and Kahneman 1974) introduced the idea that when faced with uncertain outcomes, humans tend to make decisions using specific heuristic patterns that are informed by our beliefs and experiences, rather than pursuing a purely rational analysis of possible decision outcomes. Tversky and Kahneman (1974) stated that people “rely on a limited number of heuristics which sometimes yield reasonable judgments and sometimes lead to severe and systematic errors.” A more recent review on the subject (Haselton et al. 2005) shows a variety of evolutionary psychological roots to these behaviors, recognizing that although cognitive biases depart from mathematical precision, they represent a successful human adaptation at functional problem solving. Among the many cognitive biases identified in the literature, Table 1 represents some examples of common cognitive biases that are proposed herein to be potentially relevant to remediation decision-making. Table 1 was distilled from Kahneman (2011).

In summary, intuition-based decision-making is a powerful human trait that leverages experience and heuristic short-cuts to solve problems efficiently. However, in contrast to algorithm-based quantitative decision analysis, intuition-based decisions are subject to cognitive traps that may lead to bad decisions. Nonetheless, no decision can be made on a purely quantitative basis, and both elements of decision-making are important and are involved in most decisions. Furthermore, the author claims no immunity from the cognitive biases listed in Table 1, and the reader is left to consider how some of them may come into play in their own remediation decisions.

Objectives

The overall objective of the remediation decision-making survey was to assess decision-making practices in the groundwater remediation industry, particularly the relative prevalence of quantitative decision analysis vs. gut intuition-based practices. Several specific objectives were established, as follows:

- Characterize the background and professional demographics of survey respondents, in order to provide context to the results.
were generally oriented to the aforementioned objectives.

S1 (Supporting information). A survey of remediation industry practitioners was conducted using the SurveyMonkey™ website during May and June 2015. A total of 118 respondents completed the survey which required 10 to 20 min to be completed. The full survey including responses obtained is available in Appendix S1 (Supporting information).

The survey design included a total of 34 questions that were generally oriented to the aforementioned objectives.

### Table 1
Examples of Cognitive Biases That May Be Relevant to Remediation Decision-Making

| Planning fallacy | Incorrect belief that a decision alternative is more likely to succeed exactly as planned, if accompanied by a high level of planning |
|------------------|----------------------------------------------------------------------------------------------------------------------------------|
| Confirmation bias| Excessive weighting of information that is supportive of a belief held by the individual, often reinforced by in-group or herd behavior, and often discounting of contradictory information |
| Neglect of probability | Failure to recognize, on a probabilistic basis, the numerous possible outcomes of a decision, including secondary and tertiary outcomes |
| Availability bias | Preferential weighting of information that is easier to recall or more recently or frequently heard |
| Loss aversion | Over-weighting of potential losses relative to potential gains |
| Anchoring effect | The undue weighting of the first decision alternative posed |
| Overconfidence effect | The expectation of a greater likelihood of individual success than would be predicted by a statistically representative success rate |
| Intuition bias | The instinctual human tendency to favor fast, intuitive thinking and avoid slower, more complex thinking and detailed analysis, in order to conserve cognitive resources needed for survival |

Most questions were categorical questions with radio-buttons to select an appropriate response, although the survey also included one simple math problem requiring entry of a numerical answer. An example question pertaining to the respondents’ demographics and experience is: Question 1. How many years of remediation experience do you have? \([0] (1 \text{ to } 5) (6 \text{ to } 10) (11 \text{ to } 15) (16 \text{ to } 20) (21 \text{ to } 25) (>25)\).

For questions such as this example, a quantitative response scale was used for analysis, based on the average of the range in each response option.

Many questions pertaining to respondents’ perceptions and experiences were structured as ranking selections with, for example, a number of items to be ranked using categories for how important they regard each item in a list \([\text{not important}] \text{ [less important]} \text{ [moderately important]} \text{ [very important]}; or how often they have used, done, or encountered each item in a list \([\text{almost never}] [\text{rarely}] [\text{sometimes}] [\text{usually}] [\text{almost always}]\). For these types of questions, the respondent was able to choose only one answer, and a Likert scale was used for analysis (Rea and Parker 2014).

In order to measure the response to each Likert scale question, an ordinal scale-level \((L)\) was assigned to each possible response, for example: \([\text{almost never}: L=1] [\text{rarely}: L=2] [\text{sometimes}: L=3] [\text{usually}: L=4] [\text{almost always}: L=5]\). A recognized weakness of Likert scale questions is that the ordinal scale may not always accurately represent differences in magnitude of perception regarding the question. In other words, the perceived difference between “rarely” and “sometimes” may not be the same as the perceived difference between “usually” and “almost always.” Likert scale questions are also subject to central-bias, where respondents may avoid answers at either extreme of a scale, such as “almost never” or “almost always.” Although absolute consensus does not exist in the statistics literature, recent work (Norman 2010; Rea and Parker 2014) shows that an arithmetic mean rank of the ordinal scale response is an appropriate measure of the response to Likert scale questions. The ranked mean \((R)\) of responses was therefore used for analysis, where \(R\) is given by:

\[
R = \frac{\sum_{i=1}^{n} n_i L_i}{n}
\]

(1)

where \(L\), Likert scale level; \(n_i\), number of levels; \(n = \text{number of responses at each scale-level; } n\), number of respondents answering the question.

For the Likert scale questions, the standard error \((SE)\), which represents an estimate of how far the calculated value of \(R\) is likely to be from the population mean (Rea and Parker 2014), was calculated for each response, as:

\[
SE = \sqrt{\frac{\sum_{i=1}^{n} n_i (L - R)^2}{n - 1}}
\]

(2)

Some questions asked the respondent to check off an unlimited number of items in a list, without mutual exclusivity. An example is: Question 33. Which of the following have you used as part of selection or design of a remediation approach (check all that apply)? For these questions, the percent of respondents checking each item was used for analysis.
Potential intuition bias was assessed using a simple math problem described by Kahneman (2011), which is described later. Potential overconfidence effect was evaluated by assessing respondents’ confidence in meeting their current project goals, as well as using a randomized A-B comparison of differences in an individual’s perceptions about either (A) their individual’s performance, or (B) overall industry performance. The A-B comparison included multiple Likert scale questions pertaining to experiences and outcomes on remediation projects. Responses for the A-B groups were evaluated using the ranked mean response, $R$ (Equation 1). The statistical significance of differences in responses between the A-B groups was assessed at a 95% confidence interval, using a one-way analysis of variance (ANOVA), as recommended by Rea and Parker (2014) and Norman (2010). ANOVA was carried out using JMP Pro v. 11 software (SAS Institute, Cary, North Carolina).

A number of uncertainties and biases may exist in the experimental design of any behavioral survey that introduce unexpected errors. An example is social desirability bias (Phillips and Clancy 1972), where survey respondents may tend to answer questions in a manner that will be viewed favorably by others. Another bias arises where responses are affected by the respondent trying to figure out the purpose of the survey (Orne 1962). There has been no attempt to unravel such influences on the survey results.

The representativeness of the survey sample to an appropriate population of remediation industry decision-makers is important. The survey sample was not randomized, and survey respondents were invited via several pertinent LinkedIn™ groups, and through the authors professional network. Because the survey sample was not random, the results may not be representative of all remediation practitioners. However, the respondents’ professional demographics described in the results suggest that the sample was representative of remediation decision-makers who are generally informed of and engaged in industry best practices. The $SE$ for any question is relatively insensitive to population size, and Equation 2 is applicable for any population over approximately 2000 individuals. Currently 412,000 people are employed in the United States in the waste management and remediation services subsector, NAICS 562 (U.S. BLS 2017). While that includes mostly individuals who are not in a decision-making role, there are certainly thousands of individuals involved in remediation decision-making, including consultants, regulators, academics, and project owners, as represented by the 1650 professionals in attendance at the 2014 Battelle International Conference on Remediation of Chlorinated and Recalcitrant Compounds.

### Results

**Respondent’s Professional Demographics**

The 118 survey respondents had experience as consultants (80%), contractors (14%), technology providers (4%), academics (14%), and private industry clients (25%) and government clients (11%). Many respondents had career experience in several areas, leading to a sum of greater than 100% above. A total of 95% of the respondents had a college degree in science, engineering, or math. A total of 90% of the respondents had read a scientific journal article in the past few months. The respondents represented an experienced group: A total of 47% of respondents had more than 20 years of experience, and 64% had a graduate degree. A total of 76% had been the project manager on a remediation project, 80% had led the selection of a remediation technology/strategy, and 64% had led the detailed design of a remediation system. The 118 respondents had worked on a total of 4900 sites over their careers (some of which may be duplicate). The median respondent had worked on 25 sites, 69% of respondents had worked on greater than 10 sites, and 34% had worked on greater than 50 sites.

**Experiences with Remediation Goals and Outcomes**

Selection of an overall project objective is the first major decision made on most remediation projects. The results of Question 13 (Figure 1), sorted by mean rank ($R$) (Equation 1) of the response to each item, indicated that short-term interim measures were the most common overall project objective/goals listed below has been in your experience at remediation sites that you have worked on.

| Objective | Mean Rank |
|-----------|-----------|
| Short-term interim measures to address migration or eminent risks | 3.17 |
| The overall objectives changed over time | 3.05 |
| Achieve a “monitoring only” status - with remedial actions completed | 2.91 |
| Do what you can within a certain budget | 2.70 |
| Contain or control contamination in perpetuity | 2.69 |
| Achieve status of remedy in place | 2.69 |
| Implemented remediation, although the end-point goal wasn’t clear | 2.64 |
| Do “something” | 2.48 |
| No Further Action with alternative end point | 2.25 |
| No Further Action / unrestricted clean closure to regulatory standard | 2.24 |
| I didn’t know the objective/goals | 1.57 |

**Q13: Please rank roughly how common each of the overall project objectives**

![Figure 1. Respondents ranking of experience with remediation objectives. For each item, the mean rank response is shown as well as error bars representing the standard error.](NGWA.org)
objective among the respondent’s experience, followed by a monitoring-only closure of remedial actions. NFA either with an alternative end point or to unrestricted clean closure to regulatory standard were the two least common objectives in the respondents’ collective experience. Question 13 allowed the respondent to rank how common each objective was in their experience, so no one answer was selected at the exclusivity of another. The survey did not evaluate differences in outcomes or other factors as a function of the objectives on a specific project.

The survey also evaluated the extent to which site closures had been attained, in the respondent’s experience. Rather than ask what percentage of sites a respondent had closed, the survey asked how many sites the respondent worked on (Question 10). Then, in Questions 11 and 12 asked: Question 11. At approximately how many sites in your career have you received a definitive No Further Action (NFA) closure—without monitoring requirements? and Question 12. At approximately how many sites in your career have you received a Monitoring-Only Closure—with completion of remedial actions? The results (Figure 2), indicated a similar number of NFA closures and monitoring-only closures; the median response was an approximately 10% closure rate for each type. Many respondents had not closed any sites: 33% of respondents reported no NFA Closures, and 30% reported no monitoring-only closures. Since these survey questions reflected the respondents’ overall experience, the results do not allow a correlation of project-specifics with closure outcomes.

The survey data showed that more experienced respondents reported higher closure rates, as measured by both the number of sites they had worked on (Figure 3) and their years of career experience (Figure 4). Considering the subset of respondents who had worked on more than 20 sites; less than 5% of the group had not closed any sites by either NFA or monitoring-only, and the group had a median closure rate of 15% for each type of closure (i.e., NFA and monitoring-only). Considering the effect of years of experience, the subset of respondents with less than 15 years of experience had a median 2% NFA closure rate, and those with more than 15 years of experience had a median 20% NFA closure rate. Fifty percent of respondents with less than 15 years of experience and 11% of those with more than 15 years of experience had zero NFA closures.

Perceptions and Experiences with Decision Inputs and Processes

Question 24 (Figure 5) evaluated the relative perceived importance of different inputs to the decision process on remediation technology selection and design. It should be recognized that a practitioner would prioritize decision inputs differently on a project-specific basis, and the question was intended to assess the overall perceived value of a variety of different decision inputs. Consistent with their mostly technical backgrounds, the respondents expressed that the most important decision inputs are strongly data-related, including detailed geology, contaminant properties, and a comprehensive conceptual site model (CSM). The bottom three least important decision inputs included all of the possible choices (3 of 15 choices) that related to the decision process itself. These results suggest that we tend to focus on...
on the data going into a decision more than how we make a decision.

While Question 24 assessed the perceived importance of various elements in decision-making, Question 33 (Figure 6) asked what information was used more commonly in practice. Some of the choices in Question 33 represent the data and calculations behind the decisions, while some represent the decision process or decision method itself. First, consistent with Question 24, the top two responses were data inputs to our decisions. Interestingly, the fourth most frequent (78%) of 38 choices was that the respondents had "asked someone you trust for their opinion," representing one of the foremost gut intuition-based decision methods (Kahneman 2011). A total of 60% of respondents stated that they had used a rule of thumb and 56% used intuition in practice, despite their low ranking in importance (Question 24, Figure 5). Probabilistic decision analysis was ranked low in both importance (Question 24, Figure 5) and in application (Question 34, Figure 6). Question 34 (Figure 6) also gives insight into the prevalence in use of a range of technical methods such as modeling, mass flux/discharge, and back diffusion calculations which is an emerging area of analysis.

Question 14 (Figure 7) pertained to what decision tools have been used by respondents, and generally confirms the results of Question 33. Of the 12 choices in Question 14, the five items pertaining to decision analysis were all reported as having been used by less than half of respondents. Question 14 showed that more than two times as many people had used a group collaborative process than had used NPV evaluation and probabilistic decision analysis. While the answers were not mutually exclusive, and many projects could involve both quantitative and group collaborative processes, this reflects the apparent importance of group decision-making among survey respondents. Group decision-making is recognized as a gut intuition-based process (Kahneman 2011) that can result in effective decisions, but is also an opportunity for group thinking and confirmation bias to come into play (Buchanan and O’Connell 2006).

### Q24: Please rate the importance of the following "inputs" to the decision process on remediation technology selection and design.

| Response Option                                      | Percent |
|------------------------------------------------------|---------|
| Field pilot test                                      | 90%     |
| Aquifer testing                                       | 84%     |
| Guidance document from EPA, ITRC, or similar         | 79%     |
| Asked someone you trust for their opinion             | 78%     |
| Groundwater flow modeling                             | 78%     |
| Contaminant transport modeling                        | 77%     |
| Bench-scale treatability test                         | 77%     |
| Risk assessment                                       | 73%     |
| Volume of treatment zone                              | 71%     |
| Assessed source contribution to plume                 | 63%     |
| Determine NAPL volume                                 | 63%     |
| Contaminant mass balance / distribution between phases or compartments | 62% |
| Hand-calculations of groundwater velocity or transport| 62%     |
| Vapor Intrusion evaluation                            | 62%     |
| 3-D visualization                                     | 60%     |
| Applied a rule of thumb                               | 59%     |
| Some kind of chemical partitioning calculation (soil to groundwater, soil to vapor, etc.) | 56% |
| Applied your intuition                                | 55%     |
| Vendor cut-sheet / specs                              | 53%     |
| Determine NAPL saturation                             | 53%     |
| Used stoichiometry or balanced reaction eqn           | 53%     |
| Incorporated duration of cleanup process               | 53%     |
| Mass flux/discharge                                   | 52%     |
| College textbook                                      | 51%     |
| Used a design spreadsheet tool from someone else      | 49%     |
| Vendor recommendation                                 | 48%     |
| Calculation from contaminant solubility or vapor pressure | 48% |
| Prepared a formal "calc sheet" with QA/QC check       | 47%     |
| Calculated a reaction rate                            | 45%     |
| Made a design spreadsheet tool                        | 44%     |
| Vendor-provided evaluation                            | 44%     |
| Using attenuation capacity to calculate an end-point   | 44%     |
| Determined groundwater residence time                 | 42%     |
| Chemical/material compatibility chart                  | 40%     |
| Percent contaminant treatment required (e.g. 95% mass or concentration reduction) | 38% |
| Vendor-provided raw performance data that was evaluated by you | 26% |
| Probabilistic decision analysis                       | 19%     |
| Back-diffusion calculations                           | 19%     |

Figure 5. Respondents ranking of the standard error.

### Q33. Which of the following have you used as part of selection or design of a remediation approach (check all that apply):

| Response Option                                      | Percent |
|------------------------------------------------------|---------|
| Field pilot test                                      | 90%     |
| Aquifer testing                                       | 84%     |
| Guidelines document from EPA, ITRC, or similar        | 79%     |
| Asked someone you trust for their opinion             | 78%     |
| Groundwater flow modeling                             | 78%     |
| Contaminant transport modeling                        | 77%     |
| Bench-scale treatability test                         | 77%     |
| Risk assessment                                       | 73%     |
| Volume of treatment zone                              | 71%     |
| Assessed source contribution to plume                 | 63%     |
| Determine NAPL volume                                 | 63%     |
| Contaminant mass balance / distribution between phases or compartments | 62% |
| Hand-calculations of groundwater velocity or transport| 62%     |
| Vapor Intrusion evaluation                            | 62%     |
| 3-D visualization                                     | 60%     |
| Applied a rule of thumb                               | 59%     |
| Some kind of chemical partitioning calculation (soil to groundwater, soil to vapor, etc.) | 56% |
| Applied your intuition                                | 55%     |
| Vendor cut-sheet / specs                              | 53%     |
| Determine NAPL saturation                             | 53%     |
| Used stoichiometry or balanced reaction eqn           | 53%     |
| Incorporated duration of cleanup process               | 53%     |
| Mass flux/discharge                                   | 52%     |
| College textbook                                      | 51%     |
| Used a design spreadsheet tool from someone else      | 49%     |
| Vendor recommendation                                 | 48%     |
| Calculation from contaminant solubility or vapor pressure | 48% |
| Prepared a formal "calc sheet" with QA/QC check       | 47%     |
| Calculated a reaction rate                            | 45%     |
| Made a design spreadsheet tool                        | 44%     |
| Vendor-provided evaluation                            | 44%     |
| Using attenuation capacity to calculate an end-point   | 44%     |
| Determined groundwater residence time                 | 42%     |
| Chemical/material compatibility chart                  | 40%     |
| Percent contaminant treatment required (e.g. 95% mass or concentration reduction) | 38% |
| Vendor-provided raw performance data that was evaluated by you | 26% |
| Probabilistic decision analysis                       | 19%     |
| Back-diffusion calculations                           | 19%     |

Figure 6. Ranking of results from 38 possible responses to Question 33. Gut-intuition decision elements are highlighted in green, and probabilistic decision analysis in blue.
Testing for Cognitive Bias: Overconfidence Effect
As described in the Methods section, several questions comprised a randomized A-B comparison of either (A) (viewed by 41% of respondents) “your individual remediation experience on projects you have performed,” or (B) (viewed by 59% of respondents) “your impressions of overall industry-wide remediation experience on projects performed by others”. This A-B comparison assessed differences in perceptions about an individual’s own performance vs. overall industry performance. The premise was that if overconfidence effect were present, it might be reflected in a perception of better individual performance than of overall industry performance (Kahneman 2011).

Results to Questions 17 and 18, which pertained to goals and outcomes on remediation projects, are shown in Figure 9 with separate responses shown for the two groups who answered pertaining to “your experience” or “industry/ others experience.” Qualitatively, the Likert Scale scores of respondents are remarkably similar between the A-B groups. The results of the one-way ANOVA are shown in the text field for each question on Figure 9. There is not a statistically significant difference between the groups on any of the responses, at a 95% confidence interval (critical p value = 0.05). However, it is worth noting that it was nearly statistically significant (p = 0.0576) that “industry/others” were perceived to more often make a decision to “abandon a remedial action and do something else.”

The potential for overconfidence effect was also evaluated in a different manner, by asking separately about to what extent remediation goals are generally reasonably achievable, and respondent’s confidence in meeting their specific remediation goals. Respondents were asked: Question 21. Generally speaking, how often are the typical remediation goals generally reasonably achievable? Later, respondents were asked: Question 25. Think of the ‘main’ remediation site you are currently working on. For this site, what is your current confidence that the specific established measurable remediation goal will be met by the selected remedy? The mean responses for these questions indicated that goals are considered reasonably achievable 47% of the time, while individuals were 70% confident that they would be successful in meeting the specific established measurable goal on their main current project. This result may reflect a degree of overconfidence in our individual likelihood of meeting remediation goals.

Evaluating Intuition Bias Using the Bat and the Ball Problem
The reader is encouraged to pause for a moment and answer the following survey question, then record their answer. Question 34. Here is a simple word problem for you...
to solve. A bat and ball cost $1.10. The bat costs $1 more than the ball. How much does the ball cost?

Question 34 is borrowed from a classical cognitive test used to evaluate the tendency toward intuition-based fast thinking instead of more detailed (and slower) analytical problem solving (Kahneman 2011). The seemingly obvious answer to the question is that the ball costs 10 cents, but that is incorrect. A simple double-check (0.1 + (0.1 + 1) = 1.2) shows that the answer is incorrect, and either trial and error or a simple algebraic solution will determine that the ball costs 5 cents. Ultimately, the bat and ball problem represents a test of whether an individual is satisfied with a quick and intuitively comfortable answer or if they will take more time to either check the result, or analytically determine a correct answer.

In the remediation decision-making survey, 52% of respondents answered Question 34 correctly and 48% answered incorrectly, illustrating some tendency toward fast thinking and intuition-based decision-making in the remediation industry. For comparison, Kahneman (2011) reports that more than 50% of students at Harvard, MIT, and Princeton followed their intuition and answered this question incorrectly, and at less selective institutions the rate was more than 80%. Following Kahneman (2011) and others, the survey results to Question 34 were used to separate respondents into two groups: one that answered Question 34 incorrectly (n=36) and has a greater tendency toward intuition bias and another group (n=39) that answered the problem correctly and has a lesser tendency toward intuition bias.

Table 2 summarizes a number of relevant survey results across these two groups. None of the differences between these two groups in the results in Table 2 meets the criteria for statistical significance, at a 95% confidence interval, using either a T-test statistic (critical T value = 1.96) or ANOVA (critical p value = 0.05), as appropriate. Therefore, we cannot conclude that the two groups have different results in these areas or that individuals with less intuition bias have better project outcomes, as measured by site closure rates. Nonetheless, the results are qualitatively interesting, and may tempt the reader to invoke confirmation bias in their interpretation.
The survey results reported herein only scratched the surface of how we make remediation decisions, and while some of the aspects evaluated herein may seem distinct, they are probably highly interrelated. For example, respondents frequently used a trusted opinion and group decision-making, which are recognized intuitive decision methods. However, organizational and cultural factors are likely also involved, such as obtaining buy-in and distributing accountability and risks of potential failure. With respect to risk of failure, loss aversion and overconfidence are probably interrelated. Respondents were an average of 70% confident that they would be successful in meeting the specific established measurable goals on their current main project, while responding that goals are reasonably achievable an average of 47% of the time. This may reflect some degree of overconfidence, but an equally important set of questions pertains to the remaining 30% to 53% potential for not meeting goals. The survey reported herein did not evaluate whether we accounted for the probability of failure, either high or low.

Humans have a tendency toward fast-thinking and intuition bias, and based on the answers to the bat-and-ball problem, the remediation industry is no exception. However, there was not a statistically significant difference in outcomes such as NFA closure rates between respondents exhibiting greater or lesser intuition-bias, based on the bat-and-ball problem. This result reminds us that a tendency toward intuitive decision-making is a natural human trait that may not be inherently inferior.

The survey results reported herein only scratched the surface of how we make remediation decisions, and while some of the aspects evaluated herein may seem distinct, they are probably highly interrelated. For example, respondents frequently used a trusted opinion and group decision-making, which are recognized intuitive decision methods. However, organizational and cultural factors are likely also involved, such as obtaining buy-in and distributing accountability and risks of potential failure. With respect to risk of failure, loss aversion and overconfidence are probably interrelated. Respondents were an average of 70% confident that they would be successful in meeting the specific established measurable goals on their current main project, while responding that goals are reasonably achievable an average of 47% of the time. This may reflect some degree of overconfidence, but an equally important set of questions pertains to the remaining 30% to 53% potential for not meeting goals. The survey reported herein did not evaluate whether we accounted for the probability of failure, either high or low.

Humans have a tendency toward fast-thinking and intuition bias, and based on the answers to the bat-and-ball problem, the remediation industry is no exception. However, there was not a statistically significant difference in outcomes such as NFA closure rates between respondents exhibiting greater or lesser intuition-bias, based on the bat-and-ball problem. This result reminds us that a tendency toward intuitive decision-making is a natural human trait that may not be inherently inferior.

The survey results reported herein only scratched the surface of how we make remediation decisions, and while some of the aspects evaluated herein may seem distinct, they are probably highly interrelated. For example, respondents frequently used a trusted opinion and group decision-making, which are recognized intuitive decision methods. However, organizational and cultural factors are likely also involved, such as obtaining buy-in and distributing accountability and risks of potential failure. With respect to risk of failure, loss aversion and overconfidence are probably interrelated. Respondents were an average of 70% confident that they would be successful in meeting the specific established measurable goals on their current main project, while responding that goals are reasonably achievable an average of 47% of the time. This may reflect some degree of overconfidence, but an equally important set of questions pertains to the remaining 30% to 53% potential for not meeting goals. The survey reported herein did not evaluate whether we accounted for the probability of failure, either high or low.

Humans have a tendency toward fast-thinking and intuition bias, and based on the answers to the bat-and-ball problem, the remediation industry is no exception. However, there was not a statistically significant difference in outcomes such as NFA closure rates between respondents exhibiting greater or lesser intuition-bias, based on the bat-and-ball problem. This result reminds us that a tendency toward intuitive decision-making is a natural human trait that may not be inherently inferior.
in the decision process or as a continency, or if we engaged in neglect of probability which is another known cognitive bias. There are many unanswered questions remaining about our decision-making on remediation projects.

Ultimately, a key recommendation arising from the survey results is for practitioners to be mindful of the decision-making process being applied in any situation. Intuition-based decision-making using heuristics such as rules of thumb, common sense, group decision-making, and educated guessing is a powerful human ability that leverages (and requires) prior experience to render good judgments. Intuitive decisions are especially useful in the absence of adequate time or information to support more detailed decision analysis. However, intuitive decisions are subject to cognitive biases that can lead to significant failures if decision makers are not aware of them. Even decision tools such as multicriteria ranking systems outlined in many regulatory frameworks may not function as objective analytical decision tools, if the inputs are derived using intuition-based methods subject to cognitive biases.

In complex data-based decision-making, best outcomes can be expected from decisions that combine intuitive judgment with quantitative decision analysis. Even a simple form of algorithm-based decision method can lend structure and rationality to our decisions and help avoid decision traps resulting from cognitive bias (Kahneman 2011). Probabilistic decision analysis tools, such as decision trees, are particularly valuable for decisions involving uncertain outcomes. Even for less complex decisions, these tools have the advantage of guiding practitioners to recognize low-probability potential outcomes that may have extremely negative consequences.

Acknowledgments

The remediation decision survey was originally conducted under employment by Trihydro Corporation, as part of their Technical Leadership Program training and technical communication efforts. Extremely valuable comments were provided by two anonymous reviewers. The author also thanks the 118 survey respondents for their time and contribution.

Supporting Information

Appendix S1. Survey of Remediation Decision Making Practices Including Responses Obtained.

References

Buchanan, L., and A. O’Connell. 2006. A brief history of decision making. Harvard Business Review 84, no. 1: 32–41, 132.

Crumbling, D.M., C. Groenjes, B. Lesnik, K. Lynch, J. Shockley, J. van Ee, R. Howe, L. Keith, and J. McKenna. 2001. Managing uncertainty in environmental decisions: Applying the concept of effective data to contaminated sites could reduce costs and improve cleanups. Environmental Science & Technology 35, no. 18: 404A–409A.

Favara, P. 2010. Use of probabilistic tools to improve remedial decision making. In Remediation of Chlorinated and Recalcitrant Compounds. Ninth International Conference on Remediation of Chlorinated and Recalcitrant Compounds, Monterey, CA, May, ed. K.A. Fields, and G.B. Wickramanayake. Columbus, OH: Battelle Memorial Institute.

Federal Aviation Administration (FAA). 2008. Aeronautical Decision Making. FAA-P-8740-69.

Fox, J. 2014. Instinct can beat analytical thinking. Harvard Business Review: https://hbr.org/2014/06/instinct-can-beat-analytical-thinking.

Franklin, B. 1772, 1975. To Joseph Priestley. In The Papers of Benjamin Franklin: January 1 through December 31, 1772, Vol. 19, ed. W.B. Willcox, 299–300. ISBN 0300018657. New Haven: Yale University Press.

Gladwell, M. 2005. Blink: The Power of Thinking Without Thinking. 277. New York: Little, Brown and Co.

Haselton, M.G., D. Nettle, and P.W. Andrews. 2005. The evolution of cognitive bias. In The Handbook of Evolutionary Psychology, ed. D.M. Buss, 724–746. John Wiley & Sons Inc.: Hoboken, NJ.

Kahneman, D. 2011. Thinking, Fast and Slow, 499. New York: Farrar, Straus and Giroux.

Keeney, R.L., and H. Raiffa. 1976. Decisions with Multiple Objectives. New York: Wiley.

Kiker, G.A., I. Bridges, A. Varghese, T. Seager, and I. Linkov. 2005. Application of multicriteria decision analysis in environmental decision-making. Integrated Environmental Assessment and Management 1, no. 2: 95–108.

National Research Council (NRC). 2013. Alternatives for Managing the Nation’s Complex Contaminated Groundwater Sites. Washington, DC: National Research Council, Committee on Future Options for Management in the Nation’s Subsurface Remediation Effort, Water Science and Technology Board, National Academies Press.

Norman, G. 2010. Likert scales levels of measurement and the “laws” of statistics. Advances in Health Sciences Education: Theory and Practice 15, no. 5: 625–632.

Oregon Department of Environmental Quality (Oregon DEQ). 2003. Risk-Based Decision Making for the Remediation of Petroleum-Contaminated Sites.

Orne, M.T. 1962. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. American Psychologist 17, no. 11: 776–783. https://doi.org/10.1037/h0043424

Phillips, D.L., and K. Clancy. 1972. Some effects of “social desirability” in survey studies. American Journal of Sociology 77, no. 5: 921–940.

Raiffa, H. 1968. Decision Analysis: Introductory Lectures on Choices under Uncertainty. Reading, MA: Addison-Wesley.

Rea, L., and R. Parker. 2014. Designing and Conducting Survey Research, 4th ed., 332. San Francisco, California: Jossey-Bass.

Schuyler, J., and P. Newendorp. 2013. Decision Analysis for Petroleum Exploration, 3rd ed., 588. Aurora, Colorado: Planning Press.

Stegman, A. 2014. The use of decision tree models on remediation projects for effective financial and strategic decision making. In Remediation of Chlorinated and Recalcitrant Compounds. Ninth International Conference on Remediation of Chlorinated and Recalcitrant Compounds, Monterey, CA, May, ed. H.V. Rectanus, and S.H. Rosansky. Columbus, OH: Battelle Memorial Institute.

Stemole, F.J., and J. Stemole. 2009. Economic Evaluation and Investment Decision. New York: Investment Evaluations Corporation.

Stiell, I.G. 2000. Clinical decision rules in the emergency department. Canadian Medical Association Journal 163, no. 11: 1465–1466.

Tversky, A., and D. Kahneman. 1974. Judgment under Uncertainty: Heuristics and Biases. Science 185, no. 4157: 1124–1131.
U.S. Bureau of Labor Statistics (U.S. BLS). 2017. Workforce statistics: Waste management and remediation services: NAICS 562. https://www.bls.gov/iag/tgs/iag562.htm#workforce (accessed August 2017).

U.S. Environmental Protection Agency (U.S. EPA). 2005. Office of Superfund Remediation and Technology Innovation. Decision Support Tools—Development of a Screening Matrix for 20 Specific Software Tools. Washington, DC: U.S. Environmental Protection Agency.

U.S. EPA, National Risk Management Research Laboratory. 2011. EPA Growing DASEES (Decision Analysis for a Sustainable Environment, Economy & Society)—To Aid in Making Decisions on Complex Environmental Issues. EPA 600/F/023.

United States Army (U.S. Army). 2015. Center for Army Lessons Learned, Military Decision Making Process (MDMP). Hand- book No. 15-06.

Biographical Sketch

W.S. Clayton, Ph.D., P.E., P.G., is Affiliate Faculty at Colorado School of Mines, Golden, CO, 80401; wclayton@mines.edu