Framing bias: The effect of figure presentation on seismic interpretation

Juan Alcalde[1], Clare E. Bond[1], and Charles H. Randle[1]

Abstract

Interpreters of reflection seismic data generally use images to disseminate the outcomes of their geologic interpretation work. The presentation of such interpretation images can generate unwanted biases in the perception of the observers, an effect known as "framing bias." These framing biases can enhance or reduce the confidence of the observer in the presented interpretation, independently of the quality of the seismic data or the geologic interpretation. We have tested the effect of presentation on confidence in interpretation of 761 participants of an online experiment. Experiment participants were presented with seismic images and interpretations, deliberately modified in different aspects to introduce potential framing biases. Statistical analysis of the results indicates that the image presentation had a subdued effect on participants' confidence compared with the quality of the seismic data and interpretation. The results allow us to propose recommendations to minimize biases in the observers related to the presentation of seismic interpretations: (1) interpretations should be shown with the seismic data in the background to ease comparison between the uninterpreted-interpreted data and the subsequent confidence assessments; (2) seismic data displayed in color aids in the interpretation, although the color palettes must be carefully chosen to prevent unwanted bias from common color spectrum in the observers; and (3) explicit indication of uncertainty by the interpreters in their own interpretation, which was deemed useful by the participants.

Introduction

Images are one of the most powerful tools for transferring information and knowledge (Latour, 1986). Scientists use imagery to present data and information. Such imagery often allows complex or large volumes of numerical data to be visualized (e.g., in a graph or a color plot). Geoscientists are no exception using images in presentations and scientific papers to present data and to illustrate their interpretations (Edelson et al., 1999; Libarkin and Brick, 2002; Kastens and Ishikawa, 2006; Liben and Titus, 2012). The plotting or presentation of data in science is not purely used as a method for conveying results, but often forms an essential part of the interpretation process. For example, interpretation of images of geophysical data (e.g., reflection seismic data) is the main method of subsurface geologic model creation. Uncertainty in the interpretation of geophysical data and its impact on interpretation outcome is the subject of several studies (Rankey and Mitchell, 2003; Bond et al., 2007, 2012; Polson and Curtis, 2010; Macrae et al., 2016). These uncertainties and interpretation outcomes are influenced by bias. Here, we focus on how the presentation of seismic image data and associated interpretations may affect the confidence of geologists in the interpretation presented, to address the question of whether framing bias influences confidence in interpretations.

The presentation of an image can change the perception of an observer, to the extent that the presentation can modify the information transferred. Images are crafted to show specific information to an observer. In marketing, the image designer will try to convey information to a potential buyer that will maximize sales. This could be done by linking the product to a lifestyle or highlighting a reduction in price (e.g., "20% off"). The designer's job is to play on the biases and perceptions of potential buyers to influence them to buy. The presentation of images in science is often designed to highlight specific information or elements, but it should not bias the observer in his or her own interpretation of the results. Recent examples show that commonly used color scales in scientific plots bias the perception of observers (Borkin et al., 2011; Hawkins, 2015). Presentation of seismic imagery to interpreters and those evaluating interpretations should be made with care to ensure minimal bias. However, the creation of a seismic image is

[1]University of Aberdeen, Geology and Petroleum Geology, School of Geosciences, Kings College, Aberdeen, UK. E-mail: juan.alcalde@abdn.ac.uk; clare.bond@abdn.ac.uk; charles.randle@abdn.ac.uk.

Manuscript received by the Editor 5 May 2017; revised manuscript received 19 June 2017; published ahead of production 24 August 2017; published online 04 October 2017. This paper appears in *Interpretation*, Vol. 5, No. 4 (November 2017); p. T591–T605, 7 FIGS., 5 TABLES.

http://dx.doi.org/10.1190/INT-2017-0083.1.

itself based on the assumptions of the processor who creates the image, the model used to transfer the seismic data into a seismic image is often based on an assumption of layer-cake strata, and as such interpretation is of an already-biased image (Bond, 2015). Here, we do not address these initial biases but focus on the presentation of a seismic image and the associated interpretation to an evaluator of that interpretation to understand the effect of image visualization on perception. This is essential to maximize the effectiveness of information transferred from an image (Rapp and Uttal, 2006) and to minimize bias in the evaluation of interpretations.

Tversky and Kahneman (1981) propose that the way in which a problem is framed, or presented, can affect the decisions made upon them. This effect is known as framing bias, and its influence in decision making has been widely studied in several different fields such as marketing (e.g., Malkoc and Zauberman, 2006), environmental issues (e.g., Patt and Zeckhauser, 2000), economics (e.g., Dellavigna, 2009), medicine (e.g., De Martino et al., 2006), law (Busey and Loftus, 2007), and politics (e.g., Slothuus and De Vreese, 2010). As an example, the positive framing of an attribute of a product can increase the favorable opinion of a customer about it, and therefore encourage its purchase (e.g., Levin, 1987; Zhang and Buda, 1999). Attribute framing is also present in visual representations of information. Multiple experiments, mainly in the field of marketing, show that the attention of the viewer can be intentionally drawn toward target areas by modifying the saliency of different parts of an image (i.e., making them stand out from other parts) (Gamliel and Kreiner, 2013). This can have a strong effect on the evaluation of an image (e.g., Sun et al., 2012) and modify the emotional response induced by them (e.g., Era et al., 2015).

When a geoscientist evaluates a seismic interpretation carried out by another geologist, the evaluation involves observing the seismic data and the interpretation together, to assess the quality of the data and the coherence of the interpretation to that data and to geologic rules or reasoning (e.g., Frodeman, 1995; Bond, 2015; Alcalde et al., 2017a). From this visual assessment, the geologist will form an opinion on the interpretation and how confident they are in it. Because the amount of information transferred by an image and the confidence of observers in an interpretation could be strongly affected by the presentation of the image, determining the presentation configuration that allows the most effective information exchange while minimizing bias in the observer is desirable. To test this, we conducted an online experiment in which participants were asked to assess their confidence in different seismic interpretations, composed of an uninterpreted seismic image and its interpretation. These images were modified to test the effect of different aspects of the presentation (variables) in the confidence of the observer in the interpretation from potential framing bias. The presentation variables included changes in aspects of the seismic image data (e.g., presenting the seismic image in color versus grayscale) and the interpretation (e.g., use of digital versus hand-drawn interpretations). We analyze statistically the results of the online survey to assess the variables that impact the confidence of participants in the interpretation and offer recommendations on how to present interpretations and seismic image data, based on our findings.

Experiment setup

In total, 761 participants took part in an online experiment to investigate how differences in the presentation of interpreted seismic images affected participant confidence in the interpretations. The experiment was released online on the 14 April 2016, and it was available for 42 days. The participants were geoscience students, academics and industry professionals from around the world. The online survey platform used to deliver the experiment (Survey Monkey) allowed a large number of participants to be reached. The distribution channels included geosciences mailing lists (e.g., Meeting of Young Researchers in the Earth Sciences [MYRES], British Geophysical Association [BGA], and British Sedimentological Research Group [BSRG]), social networks (e.g., Twitter, LinkedIn, and Facebook), and word-of-mouth (WOM). On the day the experiment was released, 280 participants completed the survey (representing 37% of those who completed the survey over the whole 42-day release period); 90% of the participants completed it within a week.

The experiment considered eight tests that compared two presentation variables (Figure 1). The 16 coupled variables tested are as follows:

1) "Assertive versus dubious language": Use of assertive/affirmative language in the description of the interpretation (e.g., "fault") as opposed to the use of question marks "?" indicative of uncertainty in the interpretation (e.g., fault?).
2) "Color versus grayscale seismic": Presentation of the seismic image data in color as opposed to in grayscale.
3) "Single color versus multiple color lines": Coloring of the horizons interpreted using different colors as opposed to using only black horizons.
4) "Continuous versus dashed lines": Delineation of the horizons using continuous lines as opposed to using dashed and/or dotted lines.
5) "Digital versus hand-drawn interpretation": Presentation of the interpretations using digital drawing tools as opposed to hand drawing with colored marker pens.
6) "Seismic versus blank background": Presentation of the interpretation with the seismic section in the background as opposed to presentation of the interpretation with no seismic image data behind.
7) "Thick versus thin lines": Delineation of the horizons using thick lines as opposed to using thin lines.
8) "Weak versus strong seismic contrast": Presentation of the seismic image data with a reduction of 20% in the contrast of the image, as opposed to using the

same image with the image contrast enhanced by 20%. The change in contrast was applied using the correction tool in Microsoft PowerPoint.

Each image presented consisted of the uninterpreted seismic image data plus its corresponding interpretation, presented adjacent to each other. The images were arranged into sets so that each variable was tested twice (i.e., on two different seismic images), so a total of 32 seismic images (Figure 2a) and corresponding interpretations were used. Figure 2b shows the sets of images that the participants were given in the experiment, each participant was exposed to four out of the eight possible tests. The sets were allocated randomly to participants.

The seismic data images and interpretations were sourced from six peer-reviewed papers published in four journals, to ensure the relevance of the experimental results to actual seismic imagery and interpretations used within the geologic community (references to the papers from which the images came can be found in Table 1). Seismic images with a similar appearance were preferred, to avoid bias associated with differences in the "look" of the base seismic image. As a result, the seismic images are mainly from the same acquisition survey of the North Sea/North Atlantic and were interpreted and published originally by a single research group. All the images were presented with a vertical exaggeration of approximately 1:2 vertical to horizontal ratio, to avoid differences in perception associated with variations in the vertical and horizontal scale of the images.

The participants were asked to rate their confidence in the interpretation provided in each image, from 1 (no confidence) to 10 (maximum confidence). A rating of confidence was mandatory to continue to the next image. The discrepancy in the number of responses to the different images (Table 1, column 2) respond to participants who failed to complete the full survey. The lowest number of confidence values obtained for an image was 105. Together with the confidence ratings, the participants had the possibility of leaving comments about each image (Table 2). The text comments were optional, and there were fewer responses per image than confidence ratings; the minimum number of responses to any image was 69 (Table 2, column 2).

The participants completed a questionnaire designed to elicit background knowledge in seismic interpretation, interpretation frequency, and years of posteducation experience in seismic interpretation (Figure 3). Responses to the questionnaire were anonymous. In total, 65% of the participants have a background in geology, 18% in geophysics, and 17% in other backgrounds (physics, or not answered) (Figure 3a). Most of the participants have knowledge in seismic interpretation ranging from basic (42%) to good (33%) (Figure 3b). The frequency with which participants interpret seismic data is very variable, with 42% of them interpreting data once a month or more frequently, 27% interpreting data annually, and 31% had never interpreted seismic data, or that left this question unanswered (Figure 3c). The posteducation experience of the participants ranged from zero years (i.e., students) up to one participant with 65 years of experience (Fig-

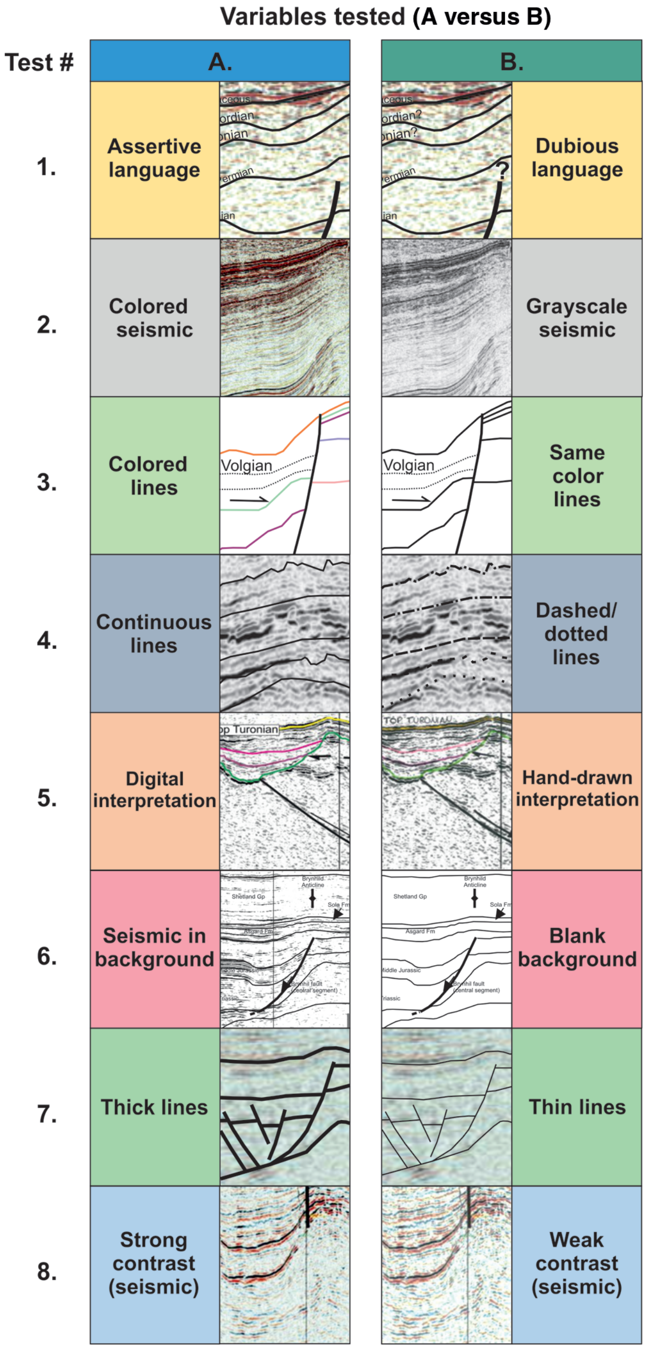

Figure 1. Example of the presentation variables tested in the online interpretation survey: (1) assertive versus dubious language; (2) color versus black and white seismic background; (3) colored versus same color horizon lines; (4) continuous versus dashed and dotted horizon lines; (5) digital versus hand-drawn interpretation; (6) seismic versus blank background; (7) thick versus thin lines; and (8) weak versus strong seismic contrast. These examples are portions of the real figures used in the experiment (references of the figures in Table 1).

ure 3d), although most of the participants have 3–15 years of posteducation experience.

The experiment as presented is available in the supplementary materials can be accessed through the following link: s1.pdf.

Data analyses

The results were first analyzed graphically with histograms and boxplots of the surveyed confidences prepared for each image. This allowed an initial assessment of the result to determine the method for further analysis, as well as to draw some initial conclusions on how the confidences and image characteristics are related. Statistical analyses were performed using a Welch's *t*-test (Welch, 1947). Welch's *t*-test was used to test the hypothesis that mean confidence had been altered by changing the presentation of the image. Welch's *t*-test was chosen over alternatives due to its ability to work with our distributions that often had unequal variances and skewed distributions (Fagerland, 2012). The result is given in the form of a *p*-value, with a $p < 0.05$ (5% of significance) taken as rejection of the null hypothesis, indicating that changing the image had influenced the confidence.

Analysis of the text responses was challenging due to the number (more than 2600 responses), the complexity (e.g., multiple sentences per comment), and the intrinsic subjectivity of the responses. We examined the comments to determine if they addressed any aspects of the interpretation, data, or the presentation of the images ("I," "D," or "P," respectively), and whether these remarks were positive (+) or negative (−) (Table 2). This count allowed us to extract quantitative information about the text comments. To avoid bias in categorizing the comments, the confidence values were obscured during manual analysis. Only clear statements were taken into account; for example, a participant's comment that stated "*Shape of normal faults*" was not considered clearly positive or negative and was not included in the count. When a single comment contained negative and positive statements, e.g., "*major reflectors are clear. The continuation of the largest faults at depth is questionable*," it was counted as an I+ and I−. Only one positive and/or negative count was added per aspect (I, D, or P) for each participant even if the comment indicated multiple positive

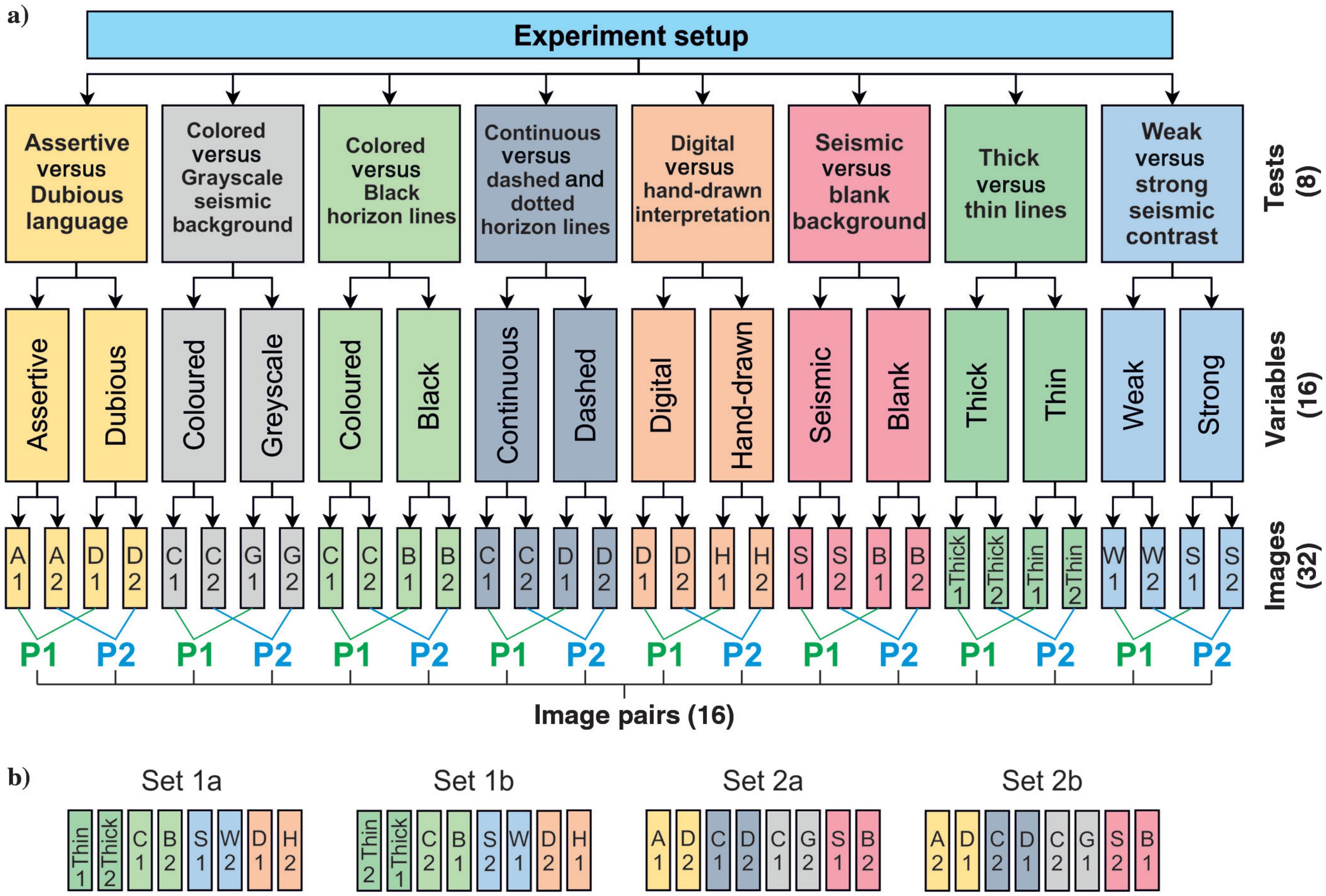

Figure 2. Experiment setup: (a) the eight presentation tests involved the comparison of two variables, each of which was tested on two different seismic images, resulting in a total of 32 images, and (b) the 32 images were divided in two pairs of conjugate sets; the sets were randomly presented to the experiment participants. Note that sets 1a and 2a are the conjugate of sets 1b and 2b, respectively, and therefore each participant was presented with alternate images of the variables of four tests. For clarity, the colors used in the boxes of each test are the same throughout the manuscript.

or negative elements for a single aspect (e.g., "*good continuity of reflectors; nice contrasts*" counted only once for D+). This counting method prevents long responses from individual participants carrying more weight than shorter responses in the global count.

Results

Confidence results

The 761 participants provided 4038 assessments of their confidence of the seismic interpretations presented. The confidence assessments include all possibilities (1–10), but most of the responses, those that fall within the first and third quartiles of the overall response distribution, lie in the confidence range of 3–7 (Figure 4). The ranges of confidence distributions were calculated for each tested image, they show a wide variety in the mean confidence (Table 1; Figure 5), with values ranging from 3.5 to 6.7. The results of Welch's *t*-test for each pair of images are presented in Table 3. The *p*-values for the seismic versus blank background test for image 1 (*p*-value = 0.04) and image 2 (*p*-value = 0.002) show that there is a statistically significant difference in mean confidence between interpretations in which a seismic background is shown and not shown behind the interpretation, with confidence being greater when a seismic background is shown. The results of the colored versus grayscale seismic background and the digital versus hand-drawn tests are less conclusive, with both having a statistically significant difference in mean confidence for one test image but not the other. This may

Table 1. Mean confidence responses for each image tested, ordered by mean confidence value, including the reference of the original figure used for each test.

Test and image	Total number of responses	Mean confidence	Reference of the modified figure
Digital versus hand drawn 1: D	157	6.69	Jackson and Larsen (2009) (Figure 4a)
Assertive versus dubious 2: D	129	6.58	Kane et al. (2010) (Figure 7)
Assertive versus dubious 1: A	154	6.57	Jackson and Larsen (2008) (Figure 7d)
Colored versus same 2: S	145	6.36	Kane et al. (2010) (Figure 7)
Weak versus strong 2: W	132	6.33	Kieft et al. (2010) (Figure 9)
Weak versus strong 1: S	135	6.26	Jackson and Larsen (2008) (Figure 7e)
Colored versus same 1: C	145	6.22	Jackson and Larsen (2008) (Figure 7b)
Seismic versus blank 2: S	104	6.20	Jackson and Larsen (2008) (Figure 4)
Digital versus hand drawn 2: H	157	6.16	Jackson et al. (2008) (Figure 6a)
Colored versus grayscale 1: C	134	6.02	Jackson et al. (2013) (Figure 8)
Seismic versus blank 1: S	129	5.84	Jackson et al. (2008) (Figure 12)
Weak versus strong 2: S	117	5.76	Kieft et al. (2010) (Figure 9)
Weak versus strong 1: W	125	5.58	Jackson and Larsen (2008) (Figure 7e)
Colored versus grayscale 2: B	117	5.57	Kane et al. (2010) (Figure 7)
Seismic versus blank 1: B	121	5.38	Jackson et al. (2008) (Figure 12)
Continuous versus dashed 2: D	120	5.34	Jackson et al. (2008) (Figure 6c)
Seismic versus blank 2: B	113	5.30	Jackson and Larsen (2008) (Figure 4)
Thick versus thin 1: thick	133	5.13	Jackson and Larsen (2009) (Figure 4b)
Colored versus grayscale 2: C	105	5.09	Kane et al. (2010) (Figure 7)
Thick versus thin 2: thin	126	5.06	Kieft et al. (2010) (Figure 8)
Continuous versus dashed 1: C	141	5.02	Jackson and Larsen (2009) (Figure 4c)
Digital versus hand drawn 2: D	122	4.97	Jackson et al. (2008) (Figure 6a)
Thick versus thin 1: thin	125	4.82	Jackson and Larsen (2009) (Figure 4b)
Digital versus hand drawn 1: H	129	4.79	Jackson and Larsen (2009) (Figure 4a)
Thick versus thin 2: thick	117	4.68	Kieft et al. (2010) (Figure 8)
Continuous versus dashed 2: D	106	4.68	Jackson et al. (2008) (Figure 6c)
Continuous versus dashed 1: C	124	4.66	Jackson and Larsen (2009) (Figure 4c)
Colored versus grayscale 1: B	123	4.39	Jackson et al. (2013) (Figure 8)
Colored versus same 1: S	128	4.06	Jackson and Larsen (2008) (Figure 7b)
Colored versus same 2: C	120	3.90	Kane et al. (2010) (Figure 7)
Assertive versus dubious 2: A	111	3.87	Kane et al. (2010) (Figure 7)
Assertive versus dubious 1: D	126	3.54	Jackson and Larsen (2008) (Figure 7d)

be caused by changes in presentation only affecting confidence in certain images. Because none of the other tests rejected the null hypothesis (i.e., had a p-value >0.05), we can say that none of the presentation changes to those interpretations had an effect on the participants' confidence in them.

The assertive/dubious language, color/grayscale seismic, colored/black colored lines, and digital/hand-drawn interpretations tests showed a difference in mean confidence across both image pairs (Figure 5; Table 1). This suggests that the differences in the base image/interpretation presented influence the confidence values assigned by the participants more than the framing aspects (e.g., assertive/dubious language, color/black and white seismic, color/same color lines, and digital/hand-drawn interpretations). The cause of this effect is likely related to differences in the apparent quality of the display of the two original seismic images presented to the participants.

Text responses

From the 2674 text comment responses, we extracted 3887 different aspects that could be related to the interpretation, data, and presentation, either in a positive or negative way, i.e., I+, D+, P+, I−, D−, and P− (Table 2). On average, one-third of these aspects

Table 2. Text responses of the different images tested ordered alphabetically.[2]

Test and image	Total responses	I+	D+	P+	I-	D-	P-	Average positive	Average negative
Assertive versus dubious 1: A	105	39	19	5	41	10	7	20.0%	18.4%
Assertive versus dubious 1: D	99	40	16	13	45	13	15	23.2%	24.6%
Assertive versus dubious 2: A	80	9	19	2	37	55	14	12.5%	44.2%
Assertive versus dubious 2: D	85	6	4	1	38	46	24	4.3%	42.4%
Colored versus grayscale 1: B	86	34	8	7	46	33	22	19.0%	39.1%
Colored versus grayscale 1: C	86	27	15	3	43	18	17	17.4%	30.2%
Colored versus grayscale 2: B	75	13	11	1	52	8	11	11.1%	31.6%
Colored versus grayscale 2: C	71	13	10	0	22	21	16	10.8%	27.7%
Colored versus same 1: C	99	39	22	1	44	33	13	20.9%	30.3%
Colored versus same 1: S	92	28	15	4	53	23	11	17.0%	31.5%
Colored versus same 2: C	79	17	0	0	49	43	12	7.2%	43.9%
Colored versus same 2: S	87	20	0	2	48	46	15	8.4%	41.8%
Continuous versus Dashed 1: C	94	19	4	4	45	44	16	9.6%	37.2%
Continuous versus dashed 1: D	89	31	6	2	55	28	16	14.6%	37.1%
Continuous versus Dashed 2: C	74	22	13	7	59	4	13	18.9%	34.2%
Continuous versus dashed 2: D	75	13	11	1	52	8	11	11.1%	31.6%
Digital versus hand drawn 1: D	110	63	19	6	48	30	2	26.7%	24.2%
Digital versus hand drawn 1: H	102	48	16	3	59	18	7	21.9%	27.5%
Digital versus hand drawn 2: D	74	12	2	2	58	12	14	7.2%	37.8%
Digital versus hand drawn 2: H	82	12	4	1	59	14	7	6.9%	32.5%
Seismic versus blank 1: B	80	18	9	0	60	5	20	11.3%	35.4%
Seismic versus blank 1: S	72	27	6	4	43	8	9	17.1%	27.8%
Seismic versus blank 2: B	75	17	6	1	34	31	12	10.7%	34.2%
Seismic versus blank 2: S	69	34	9	5	27	20	19	23.2%	31.9%
Thick versus thin 1: thick	89	21	2	1	60	25	13	9.0%	36.7%
Thick versus thin 1: thin	81	14	5	1	60	21	8	8.2%	36.6%
Thick versus thin 2: thick	70	22	1	1	28	47	8	11.4%	39.5%
Thick versus thin 2: thin	77	26	4	3	37	48	4	14.3%	38.5%
Weak versus strong 1: S	93	36	20	2	55	20	5	20.8%	28.7%
Weak versus strong 1: W	77	37	15	5	43	14	5	24.7%	26.8%
Weak versus strong 2: S	69	28	19	0	36	23	4	22.7%	30.4%
Weak versus strong 2: W	78	23	27	0	48	23	8	21.4%	33.8%

[2]Note: The amount of text responses containing at least one positive comment about the interpretation, data, or presentation are denoted by I+, D+, and P+, respectively; similarly, the amount of text responses containing at least one negative comment on the interpretation, data, or presentation are denoted by I−, D−, and P−, respectively.

referred to negative aspects of the images and only one-third referred to positive comments (2654 negative versus 1233 positive comments). Multiple negative comments within the same response came together more often than positive comments (i.e., more than one negative comment in the same response in 24.4% of the cases, 17.1% multiple positive comments). Participants made negative comments about the interpretation (I–) in more than half of the images (55.8% on average), followed by data D– (29.9%), and positive comments on the interpretation I+ (29.7%) (Figure 6). Presentation aspects of the images (either positive or negative) were commented on the least (3.2% and 14.4% for P+ and P–, respectively).

We calculated the correlation between the confidence assessments and the comments for all the images (Figure 6). On average, there is a strong correlation between the negative and the positive comments and the confidence assessment values given (coefficient of determination "ρ" of 0.82 for the negative comments and −0.81 for the positive comments, Figure 6a). These values indicate that participants with greater confidence in an interpretation made more positive comments on that interpretation than those that gave lower confidences and vice versa.

Specific comments on interpretation and data (Figure 6b and 6c) show an overall correlation in positive and negative comments with respect to the confidence values given, but with certain nuances. Within the positive comments, positive comments on interpretation (I+) present the highest correlation coefficient with respect to confidence ($\rho = 0.81$), whereas positive data and presentation comments (D+ and P+) present similar lower coefficients ($\rho = 0.49$ and 0.42, respectively). The negative comment correlations present differences, with a relatively constant amount of I– comments at all confidence values ($\rho = -0.24$), compared with the higher correlations of D– and P– ($\rho = -0.51$ and −0.38, respectively). In general, comments on presentation show the lowest correlation with confidence (Figure 6d), but all the correlations calculated (i.e., generally and for the different aspects, Figure 7) show statistical significance (p-values <0.05, except I– with a p-value <0.2).

Discussion

Interpretation superposition

Of the eight tests conducted in this experiment, the presence or absence of the seismic image data in the background was the presentation-related factor that mostly affected the confidence of the participants: not showing the seismic data behind the interpretation decreased confidence for both the seismic images tested (Table 3 and Figure 5). This reduction in confidence was emphasized by some of the participants in their text response (e.g., "*I don't like interpretation separated from the data*" with a confidence assessment of four). This effect in confidence reduction could be related to three factors.

1) Not showing the interpretation overlain on the seismic image creates a heavier workload for the brain because it tries to tie elements between the two separate seismic images. When observing an image, the brain performs two processes: a perceptual process that extracts information about the input and a de-

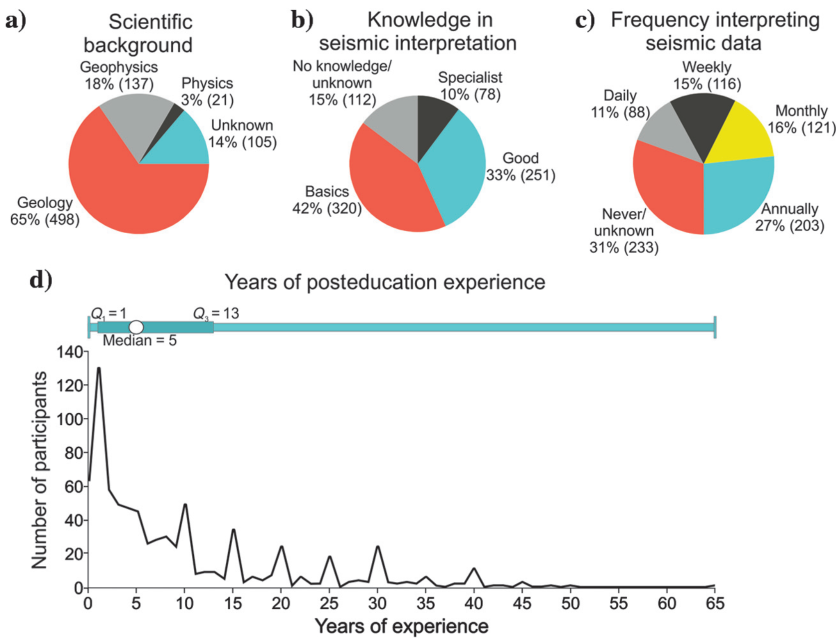

Figure 3. Statistics based on the survey participant responses about (a) their scientific background, (b) their knowledge in seismic interpretation, (c) the frequency with which they interpret seismic data, and (d) years of posteducation experience in geosciences. The peaks observed in (c) are present in the data and are associated with participants rounding up their years of experience.

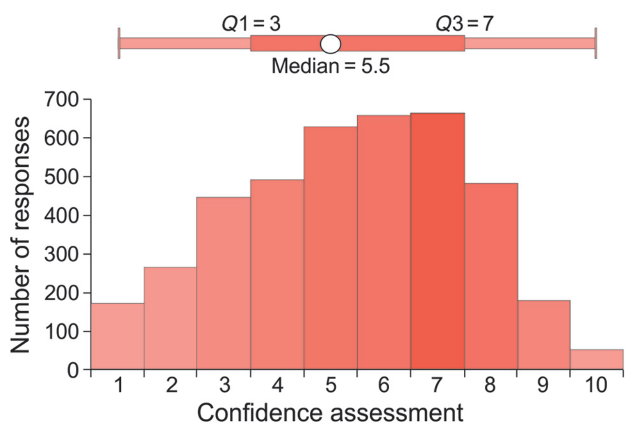

Figure 4. Histogram of confidence assessment values given by the participants in the online experiment.

cision process that evaluates the information and prepares a response (VanRullen and Thorpe, 2001). In our experiment, the participant evaluates the two images comparing the seismic image with the separate interpretation. The evaluation considers if the interpretation matches or fits the data and consequently if the interpretation is plausible or not. The brain needs to process the image information and hold onto the portions relevant for the comparison (Wedel and Pieters, 2008). In the absence of the seismic image in the background, the evaluator has to switch visually from the interpretation to the seismic image, increasing the workload of the brain, and obstructing or impeding evaluation of the images. A proportion of the brains capacity is also spent on spatially referencing the images because there is no frame of reference between the uninterpreted and the interpreted images beyond the image edge. Evaluation of an interpretation involves observing the general fit of the interpretation with the data, which includes focusing on areas of higher difficulty, perhaps where the seismic image is less clear or where the geology imaged is more complex. These difficulties are captured in the comments of the participants, e.g., "*hard to relate the interpreted horizons on the blank image to the actual seismic*" or "*Lack of seismic behind interpretation diagram makes it hard to 'see' what has been relied on for the interpretation.*" These two comments were associated with confidence assessments of two and five, respectively.

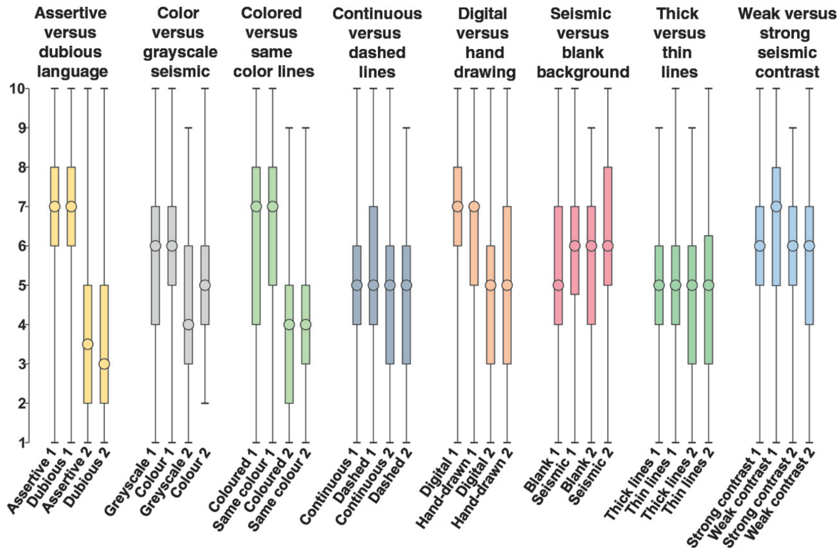

Figure 5. The box plots show the confidence value distributions of the 32 different seismic images that comprised the eight tests. The box represents the extent of the interquartile range (first and third quartile) with the position of the median marked with a circle.

Table 3. Table of *t*-test results. The results, for the eight tests, are given for each seismic image (1 or 2) as a *p*-value; *p*-values below 0.05 indicate a significant influence on confidence outcome.

Test	Image number (image)	*p*-Value
Assertive versus dubious language	1	0.97
	2	0.17
Color versus grayscale seismic	1	0.08
	2[3]	0.0038
Colored versus same color lines	1	0.58
	2	0.54
Continuous versus dashed lines	1	0.18
	2	0.95
Digital versus hand-drawn interpretation	1[3]	0.0096
	2	0.53
Seismic versus blank background	1[3]	0.04
	2[3]	0.002
Thick versus thin lines	1	0.79
	2	0.59
Weak versus strong seismic contrast	1	0.76
	2	0.50

[3]Images with *p*-value <0.05.

2) Presentation of the interpretation and seismic image separately is counterintuitive because it does not correspond to normal workflow practices. Seismic interpretation is learned and practiced, either on paper or a computer screen, by "drawing" over the seismic image. The seismic image data are always behind the interpretation while it is being carried out. This practice allows continuous analysis between the seismic image data and the interpretation. The lack of familiarity in evaluating interpretations separate from the seismic image makes the task more difficult.
3) Display of the interpretation separate to the seismic image data can create an element of suspicion that may undermine confidence in the interpretation. If the cognitive task of correlating the seismic image with the interpretation is too testing, the evaluator of the interpretation maybe mistrustful as to whether the interpretation matches the data at all. One participant commented on this: "*This looks like a reasonable interpretation, but was the profile deliberately removed from the background to hide some interpretation problems?*" but gave the interpretation a relatively high confidence value of eight. The extent of any such suspicion is likely related to the previous two points.

A combination of the aforementioned elements may not just lower confidence in the interpretation itself, but low confidence scores may be given due to the annoyance of, and or difficulty for the participants induced by the presentation of the interpretation and seismic images separately. We cannot account for this possibility in our analysis of the confidence results.

Assessment of how two related images are compared and assessed is described in the work of Gleicher et al. (2011). Gleicher et al. (2011) describe multiple comparative approaches for the visualization of related images as a mixture of three elements: juxtaposition, superposition, and explicit encoding (i.e., visual encoding of the relationships between the objects, such as colors or shapes). The seismic versus blank background test compares two of these approaches: (1) juxtaposition and superposition when the seismic image is displayed in one image with the seismic image overlain by the interpretation in an adjacent image, and (2) juxtaposition when the seismic image is shown in one image with the interpretation shown next to it on a blank background (i.e., no seismic image behind). Superposition and juxtaposition used in combination are thought to benefit transmission of information sought in the visual comparison of multiple images. It is therefore likely to enhance confidence in interpretations presented in this manner. Superposition and juxtaposition can also be helpful in the interpretation process itself, as suggested by Paton and Henderson (2015). Future work should address how the lack of a reference seismic image (i.e., superposition only) affects confidence in the interpretation. We would predict a reduction in confidence compared with the use of superposition and juxtaposition together. This is because the assessor can see the data presented in the seismic image unobscured by the interpretation, as well as the interpretation placed in context on the data, allowing an open assessment of the data and the interpretation. Given these observations, publication of uninterpreted seismic images alongside interpretations of the seismic image (i.e., juxtaposition and superposition) is recommended. The combination of juxtaposition and superposition with explicit encoding and their relationship with confidence should also be explored.

Seismic colors

For seismic image 2 in the test of color versus grayscale seismic, there is a statistical reduction in confidence of the participants in the interpretation when a grayscale seismic background is used (Table 3; Figure 5). Color, in combination with other visual factors such as intensity, hue, or perspective, determines the saliency of an object (the capacity to draw the attention of the viewer to the object). These effects are described in the context of the interpretation of seismic images by Froner et al. (2013). The human eye is capable of distinguishing several million colors, but only a few hundred shades of gray (Judd and Wyszecki, 1975). Therefore, interpreters can obtain more information from a seismic image when colored with an appropriate color palette, instead of using a grayscale palette (Chopra and Marfurt, 2005; Henderson et al., 2008; Froner et al., 2013). Moreover, the human eyes' sensitivity to color intensity varies across the color scale in a nonlinear fashion (Froner et al., 2013; Paton and Henderson, 2015). Hence, the color scale used in the seismic images can also be a source of bias for interpretation (Donnelly et al., 2006). Presentation of the seismic image in color was not mentioned in the participants' comments. However, the confidence values given were reduced when the interpretation was presented on a grayscale seismic. Only one participant commented that a colored seismic background could have helped in their assessment of the interpretation, although this comment belonged to a different test "*Doesn't look like the lower pick on SU2 is consistent across the section — having the section in colour behind the interpretation could help with this*" with a confidence value of eight in the continuous versus dashed lines test.

Given the ability to distinguish subtleties in color, beyond those presented in a grayscale, the use of color palettes for seismic displays would seem to be good practice. However, color palettes should be applied

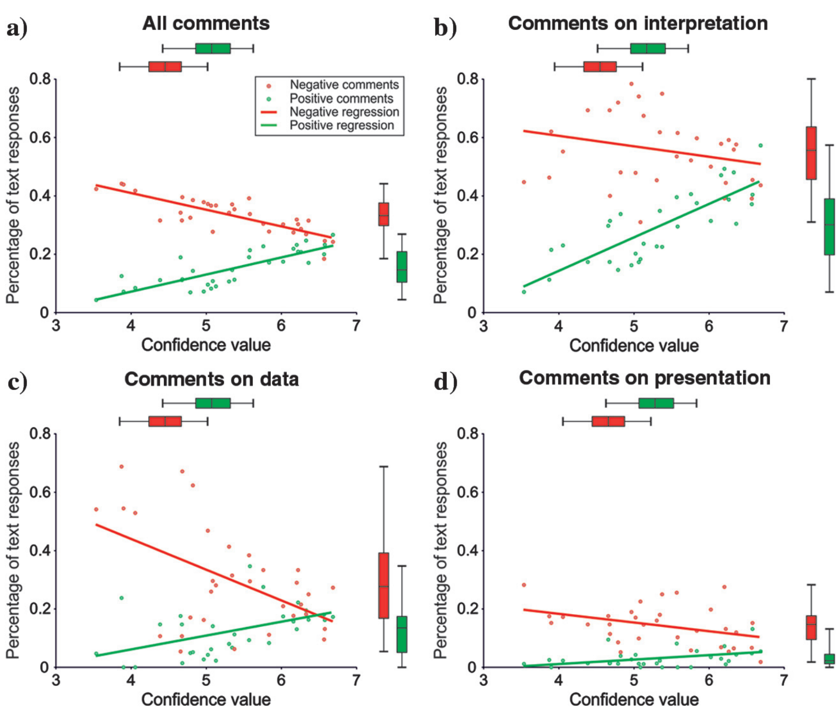

Figure 6. Correlation between percentage of positive (green) and negative (red) comments of each image versus their value of confidence in the seismic interpretation in: (a) all text comments, (b) only comments regarding the interpretation quality, (c) only comments regarding the data quality, and (d) only comments regarding the presentation quality. The linear regressions are shown to illustrate the general trends.

with caution, because some of them can produce unwanted biases, enhancing some elements while making other differences imperceptible (Donnelly et al., 2006; Welland et al., 2006). Rainbow color palettes in particular are the object of a relatively recent campaign against their use ("*#endtherainbow*"; Hawkins, 2015) because of their potential to distort the perception of the data (e.g., Borkin et al., 2011) and because they are unsuitable for the color blind. Although we have not tested different color palettes here, careful choice of color palettes and color scales should be made to aid interpretation of seismic images while diminishing the potential production of biases in interpretation (Donnelly et al., 2006; Henderson et al., 2008; Niccoli, 2014). This statement is also relevant to those evaluating interpretations of seismic image data, and further work on color scales used in seismic interpretation and evaluation would be useful.

Disassociation of confidence in interpretation and its framing

There is evidence in the experiment data that in 75% of the seismic images, participants were able to disassociate the presentation aspects (the framing) from the interpretation quality. The participants were able to see beyond, and not be influenced by, our attempts to bias the confidence ratings given, by framing the interpretations differently. The framing aspects would likely have different influences on participants' confidence ratings when compared with each other as well as influencing individual participants to different extents. Here, we discuss some examples that highlight the influence of certain framing aspects and the ability to disassociate from them.

The digital versus hand-drawn interpretation for seismic image 1, together with the two tests previously discussed (color versus grayscale seismic and seismic versus blank background), had an effect on the confidence values given by participants to the interpretation (Table 3; Figure 5). Not unexpectedly, the use of a hand-drawn interpretation for seismic image 1 reduced the confidence values given by the participants. Comments show that some deemed them as unprofessional, "*It looks like a first attempt. Confidence would be higher if it was to professional finishing standards (i.e., digitised)*" with a confidence value of four given. However, for the second seismic image tested (seismic image 2) the confidence range is not statistically different. In fact, confidence values were higher for the hand-drawn interpretation (Figure 5), and the digital interpretation of seismic image 2 had 10% more negative comments on average than the other three images in the test, including the hand-drawn interpretation for seismic image 2. Therefore, the presentation of digital interpretations does not guarantee an increase in confidence of the participants (e.g., another participant wrote: "*good data quality [be careful that block diagram and professional geologic section does not 'increase confidence']*") with a confidence value of six given. These comments show that this participant was aware of the potential for the framing of the interpretation, from its hand-drawn nature, to influence confidence in the interpretation. It is therefore assumed that in their confidence rating they attempted to disassociate the framing from the quality of the interpretation.

The lack of a consistent correlation between confidence ratings and the changes in the presentation aspects tested (Table 3), suggests that participants were

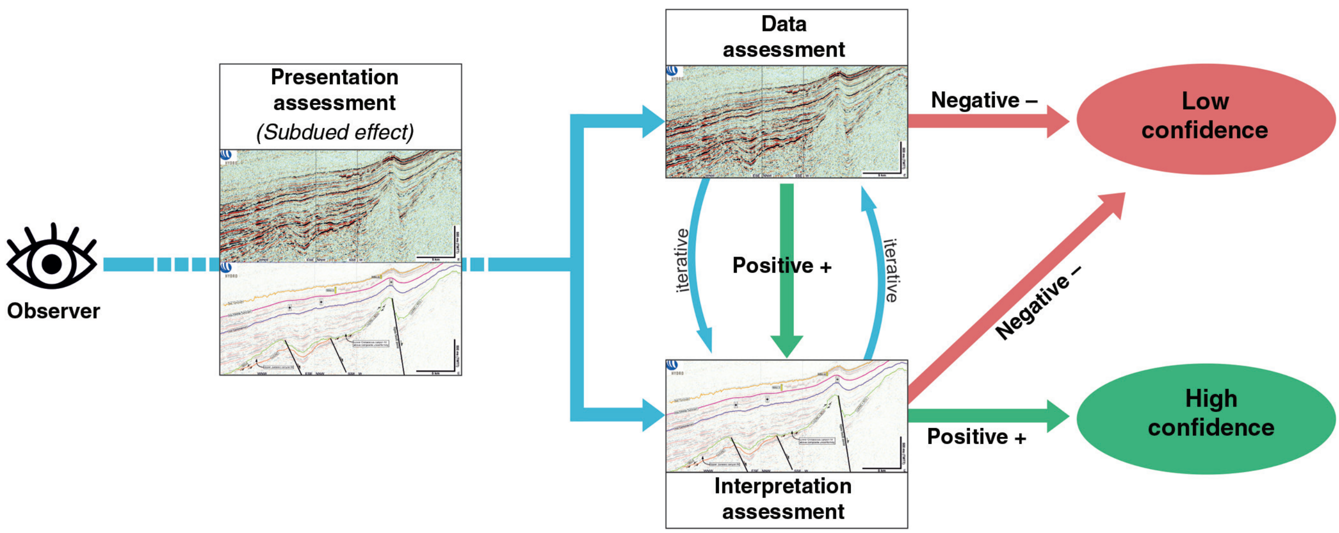

Figure 7. Proposed cognitive assessment process used by the participants based on the confidence assessment values and text comments. The participants observe the images to assess their confidence in the interpretation. The image presentation has a small effect on their confidence, so interpreters jump into the assessment of the data quality and the interpretation. This assessment is carried out by checking the data and the corresponding interpretation in an iterative way. The negative assessments of the data quality produces low confidence in the interpretation independently of the interpretation. However, the positive data assessments do not necessary lead to high confidence; it depends on assessment of the interpretation.

either disengaged from the presentation or able to disassociate from it, focusing mostly on the seismic image data quality and the interpretations. The results of the comments analysis support this argument, because presentation is commented on less than the technical aspects of the interpretation and the quality of the data (466 comments about presentation from a total of 3887) (Table 2; Figure 6). Most of the tested framing aspects had no apparent effect on the confidence of the participants in the interpretations. Therefore, in most cases, the participants were not remarkably influenced by the presentation or framing of the interpretation and concentrated on the technical geophysical and geologic aspects of the images. This is an interesting result, because there are multiple examples in which framing has a determinant effect in the decision-making process in geosciences (e.g., Chadwick, 1976; Paton and Henderson, 2015; Alcalde et al., 2017b). In our experiment, the framing of the interpretation had a subdued effect on the confidence assessment of participants. This leads to the question of what aspects of the images presented drew the attention of and influenced most the confidence of the participants in the interpretations.

Focus of the participants

Presentation of the seismic image had little influence on the confidence of the participants, but image quality was identified as an important factor. Of the three statistically significant results, two of them, color versus grayscale seismic and seismic versus blank background, include aspects of presentation of the seismic image data. The negative comments about seismic image data quality have the highest negative correlation with confidence ratings of the three types (interpretation, presentation, and data) of comments analyzed (Figure 6). The positive comments, however, present a lower positive correlation coefficient, indicating that good data are not a guarantee of high confidence in the interpretation.

Based on these findings, we propose a cognitive process for assessment of interpretations of seismic image data, which involves three factors: presentation, data, and interpretation (Figure 7). The process begins with participants observing the presented images of the seismic image data and interpretation to assess their confidence in the interpretation. Presentation (framing) of the seismic image and interpretation seems to influence participants' confidence less than the data or the interpretation, so interpreters "see through" the presentation to focus on the data and the interpretation. Participants check the quality of the data and compare the seismic image with the interpretation, to assess the quality of the interpretation. This process takes place iteratively: there is a constant movement from the data to the interpretation. If the seismic image data quality is bad, either due to a poor-quality seismic image, or perhaps the absence of a well, or other controls, further interpretation assessment becomes redundant and confidence is low. Higher confidences are likely achieved with good seismic image data that the interpretation ties to.

Explicit highlighting of uncertainty in interpretations

Seismic interpretation requires effort by the interpreter to create an interpretation of the geology, despite uncertainty in the data. Interpreters apply their skills (i.e., knowledge and understanding) to produce a consistent geologic solution that satisfies the data and their own expectations (Frodeman, 1995; Rankey and Mitchell, 2003; Bond et al., 2011). In other words, the resulting interpretation can be biased by the interpreter and their prior knowledge, which in turn may become a source of bias for the final user of the interpretation. Therefore, communicating the potential subjective uncertainty associated with an interpretation is crucial to its comprehension, although this is commonly neglected in geoscience (Bond, 2015).

The assertive versus dubious test was designed to determine if the use of question marks "?" spread over the interpretation image would impact the confidence of participants in an interpretation. The underlying idea is that an interpreter may use question marks to indicate areas in which they are less certain in their interpretation. Conversely, areas without question marks indicate a greater certainty in the interpretation. The *t*-test analysis did not reveal any statistically significant difference between interpretations using assertive or dubious language (Table 3). However, the comments left by the participants in this and other tests revealed some issues related to the perception of confidence of the interpreter in their interpretation. In spite of the lack of a statistical effect resulting from the presence of question marks in the confidence values given by participants, 21 participants praised the use of question marks to represent uncertainty, with a further eight indicating that interpretations should have included question marks or dashed lines to highlight areas of uncertainty in an interpretation. Only in three cases did the question marks reduce the confidence of the participants in the interpretations according to their comments, so in general, the explicit expression of uncertainty via question marks was perceived positively.

The words "uncertain" and "uncertainty" together appear 97 times throughout the text responses. In most cases, the explicit expression of uncertainty was praised by participants, e.g., "*on the plus side, a lot of uncertainty is expressed by the interpreter, by scattering '?' all over the interpretation — so I have confidence in its uncertainty!*" with a confidence value of five given. Furthermore, the lack of uncertainty assessment was sometimes perceived negatively as overconfidence or overinterpretation, for example "*seems over-interpreted given the relatively poor quality of the raw data and uncertainty that entails,*" with a confidence value of four given.

There is an increasing interest in the assessment and communication of uncertainty in structural geology and seismic interpretation (e.g., Bond et al., 2007; Freeman et al., 2010; Rowbotham et al., 2010; Bond, 2015; Chellingsworth et al., 2015; Richards et al., 2015). However, assessment of uncertainty of an interpretation is commonly avoided by interpreters, who perhaps are prompted to provide unequivocal results in talks, papers, or other forms of communication. This issue is boosted by the current publication system, in which scientists are encouraged to present single, deterministic solutions (Bond, 2015). This may lead to a wrong perception of certainty in interpretations and thus may be detrimental to the accurate transfer of knowledge between an interpreter and the end user. Our findings suggest that communicating uncertainties can increase confidence while presenting a better description of an interpretation, and therefore it should be encouraged.

Participants' feedback

Participants used the text response section mainly to complain about aspects of the interpretation, data, or the presentation of the images (Figure 6). Complaint behavior is documented and analyzed in the marketing field, in which individual feedback in consumer satisfaction (CS) (i.e., electronic or offline WOM) plays a major role in corporate's decisions (Lau and Ng, 2001; Henning-Thurau et al., 2004). Anderson (1998) studies CS through WOM. The author found that customers using WOM to express their satisfaction or dissatisfaction, do so to mainly express high satisfaction or dissatisfaction, thus forming a U-shaped distribution in the CS-WOM relationship. Furthermore, Anderson (1998) observes that highly dissatisfied customers used WOM mechanisms in a greater proportion than highly satisfied customers, reflected in an asymmetric U-shaped CS-WOM relationship. We can identify parallelisms between these customer satisfaction surveys and our confidence in interpretation experiment. The trends of the confidence values versus comment responses in the three groups of comments tested (i.e., interpretation, data, and presentation) show positive and negative correlations for positive and negative comments, respectively (Figure 6). Furthermore, the negative comments outnumber the positive comments in almost all the tests, even in cases in which confidence assessment was high (Table 2).

The text comments gathered in our experiment were a valuable source of information about the participants' interests and concerns in their assessment of the interpretations. The confidence values provided a good overall picture of the participants confidence in the interpretations, with the comments providing detailed information about the aspects of the images shown that the participants used in their confidence assessments. In a fuller study, this information could be used to determine data collection or processing strategies and to compare responses with different interpretations.

Experiment setup and future work

In this experiment, the number of images used in each test (2) is low compared with image-comparison experiments published in the field of vision and cognition (e.g., 54 images in Simons et al., 2000; 24 images in Čadík, 2008). Our focus was to test if framing bias influenced seismic image interpretation per se and we chose to test a wide range of variables. To minimize bias associated with "the look" of particular images we used images with a similar appearance. In future studies, testing the variables presented here over a greater number of images and image looks would increase confidence in the results.

Conclusions and recommendations

Our results suggest that specific variables in the presentation (the framing) of interpretations of seismic image data have a subdued effect on the confidence of participants in the interpretations presented. The data quality and the geologic interpretation were the primary focus of participants in determining their confidence in the interpretations on despite the framing biases that we introduced.

Based on the results of our experiment, we make three recommendations for the presentation of seismic interpretations that could help to prevent potential unwanted biases that may influence end-user assessment of interpretations:

1) Interpretations should always include the seismic image in the background with an uninterpreted seismic image next to them (superposition + juxtaposition).
2) The background seismic image should be presented in color if possible (preferably using the coloring used for the interpretation), although selection of the color palette should be chosen to avoid unwanted biases during interpretation, due to misrepresentation of the seismic image data found in commonly used color palettes.
3) The users of seismic image interpretations find indications of uncertainty in their interpretation by the interpreter useful. Such uncertainty indications can be included in the image itself or in supporting text.

Acknowledgments

The authors thank all the participants in the survey and those who helped to distribute it. We thank C. Jackson and coauthors for allowing the use of their published images in this experiment. J. Alcalde is funded by NERC grant NE/M007251/1 on interpretational uncertainty.

References

Alcalde, J., C. E. Bond, G. Johnson, R. W. H. Butler, M. A. Cooper, and J. F. Ellis, 2017a, The importance of structural model availability on seismic interpretation: Journal of Structural Geology, **97**, 161–171, doi: 10.1016/j.jsg.2017.03.003.

Alcalde, J., C. E. Bond, G. Johnson, J. F. Ellis, and R. W. H. Butler, 2017b, Impact of seismic image quality on fault interpretation uncertainty: GSA Today, **27**, 4–10.

Anderson, E. W., 1998, Customer satisfaction and word of mouth: Journal of Service Research, **1**, 5–17, doi: 10.1177/109467059800100102.

Bond, C. E., 2015, Uncertainty in structural interpretation: Lessons to be learnt: Journal of Structural Geology, **74**, 185–200, doi: 10.1016/j.jsg.2015.03.003.

Bond, C. E., A. D. Gibbs, Z. K. Shipton, and S. Jones, 2007, What do you think this is? "Conceptual uncertainty" in geoscience interpretation: GSA Today, **17**, 4–10, doi: 10.1130/GSAT01711A.1.

Bond, C. E., R. J. Lunn, Z. K. Shipton, and A. D. Lunn, 2012, What makes an expert effective at interpreting seismic images?: Geology, **40**, 75–78, doi: 10.1130/G32375.1.

Bond, C. E., C. Philo, and Z. K. Shipton, 2011, When there isn't a right answer: Interpretation and reasoning, key skills for twenty-first century geoscience: International Journal of Science Education, **33**, 629–652, doi: 10.1080/09500691003660364.

Borkin, M. A., K. Z. Gajos, A. Peters, D. Mitsouras, S. Melchionna, F. J. Rybicki, C. L. Feldman, and H. Pfister, 2011, Evaluation of artery visualizations for heart disease diagnosis: IEEE Transactions on Visualization and Computer Graphics, **17**, 2479–2488, doi: 10.1109/TVCG.2011.192.

Busey, T. A., and G. R. Loftus, 2007, Cognitive science and the law: Trends in Cognitive Sciences, **11**, 111–117, doi: 10.1016/j.tics.2006.12.004.

Ĉadík, M., 2008, Perceptual evaluation of color-to-grayscale image conversions: Computer Graphics Forum, **27**, 1745–1754.

Chadwick, P. K., 1976, Visual illusions in geology: Nature, **260**, 397–401, doi: 10.1038/260397a0.

Chellingsworth, L., M. Bentley, and T. Wynn, 2015, Human factors in seismic uncertainty — Restoring a realistic uncertainty range: Interpretation, **3**, no. 2, SQ21–SQ32, doi: 10.1190/INT-2014-0203.1.

Chopra, S., and K. J. Marfurt, 2005, Seismic attributes — A historical perspective: Geophysics, **70**, no. 5, 3SO–28SO, doi: 10.1190/1.2098670.

Dellavigna, S., 2009, Psychology and economics: Evidence from the field: Journal of Economic Literature, **47**, 315–372, doi: 10.1257/jel.47.2.315.

De Martino, B., D. Kumaran, B. Seymour, and R. J. Dolan, 2006, Frames, biases and rational decision-making in the human brain: Science, **313**, 684–687, doi: 10.1126/science.1128356.

Donnelly, N., K. R. Cave, M. Welland, and T. Menneer, 2006, Breast screening, chicken sexing and the search for oil: Challenges for visual cognition: Geological Society, London, Special publication 254, 43–55.

Edelson, D. C., M. Brown, D. N. Gordin, and D. A. Griffin, 1999, Making visualization accessible to students: GSA Today, **9**, 8–10.

Era, V., M. Candidi, and S. M. Aglioti, 2015, Subliminal presentation of emotionally negative vs positive primes increases the perceived beauty of target stimuli: Experimental Brain Research, **233**, 3271–3281, doi: 10.1007/s00221-015-4395-5.

Fagerland, M. W., 2012, T-tests, non-parametric tests, and large studies — A paradox of statistical practice?: BMC Medical Research Methodology, **12**, 78, doi: 10.1186/1471-2288-12-78.

Freeman, B., P. J. Boult, G. Yielding, and S. Menpes, 2010, Using empirical geological rules to reduce structural uncertainty in seismic interpretation of faults: Journal of Structural Geology, **32**, 1668–1676, doi: 10.1016/j.jsg.2009.11.001.

Frodeman, R., 1995, Geological reasoning: Geology as an interpretive and historical science: Geological Society of America Bulletin, **107**, 960–968, doi: 10.1130/0016-7606(1995)107<0960:GRGAAI>2.3.CO;2.

Froner, B., S. J. Purves, J. Lowell, and J. Henderson, 2013, Perception of visual information: The role of colour in seismic interpretation: First Break, **31**, 29–34, doi: 10.3997/1365-2397.2013010.

Gamliel, E., and H. Kreiner, 2013, Is a picture worth a thousand words? The interaction of visual display and attribute representation in attenuating framing bias: Judgment and Decision Making, **8**, 482–491.

Gleicher, M., D. Albers, R. Walker, I. Jusufi, C. D. Hansen, and J. C. Roberts, 2011, Visual comparison for information visualization: Information Visualization, **10**, 289–309, doi: 10.1177/1473871611416549.

Hawkins, E., 2015, Graphics: Scrap rainbow colour scales: Nature, **519**, 291, doi: 10.1038/519291d.

Henderson, J., S. J. Purves, F. Fisher, and C. Leppard, 2008, Delineation of geological elements from RGB color blending of seismic attribute volumes: The Leading Edge, **27**, 342–350, doi: 10.1190/1.2896625.

Hennig-Thurau, T., K. P. Gwinner, G. Walsh, and D. D. Gremler, 2004, Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the internet?: Journal of Interactive Marketing, **18**, 38–52, doi: 10.1002/dir.10073.

Jackson, C. A. L., G. P. Barber, and O. J. Martinsen, 2008, Submarine slope morphology as a control on the development of sand-rich turbidite depositional systems: 3D seismic analysis of the Kyrre Fm (Upper Cretaceous), Måløy Slope, offshore Norway: Marine and Petroleum Geology, **25**, 663–680, doi: 10.1016/j.marpetgeo.2007.12.007.

Jackson, C. A. L., S. T. Chua, R. E. Bell, and C. Magee, 2013, Structural style and early stage growth of inversion structures: 3D seismic insights from the Egersund Basin, offshore Norway: Journal of Structural Geology, **46**, 167–185, doi: 10.1016/j.jsg.2012.09.005.

Jackson, C. A. L., and E. Larsen, 2008, Temporal constraints on basin inversion provided by 3D seismic and well data: A case study from the South Viking Graben, offshore Norway: Basin Research, **20**, 397–417, doi: 10.1111/j.1365-2117.2008.00359.x.

Jackson, C. A. L., and E. Larsen, 2009, Temporal and spatial development of a gravity-driven normal fault array:

Middle-Upper Jurassic, South Viking Graben, northern North Sea: Journal of Structural Geology, **31**, 388–402, doi: 10.1016/j.jsg.2009.01.007.

Judd, D. B., and G. Wyszecki, 1975, Color in business, science, and industry, 3rd ed.: Wiley-Interscience, Wiley Series in Pure and Applied Optics.

Kane, K. E., C. A. L. Jackson, and E. Larsen, 2010, Normal fault growth and fault-related folding in a salt-influenced rift basin: South Viking Graben, offshore Norway: Journal of Structural Geology, **32**, 490–506, doi: 10.1016/j.jsg.2010.02.005.

Kastens, K. A., and T. Ishikawa, 2006, Spatial thinking in the geosciences and cognitive sciences: A cross-disciplinary look at the intersection of the two fields: Geological Society of America Special Papers, **413**, 53–76.

Kieft, R. L., C. A. L. Jackson, G. J. Hampson, and E. Larsen, 2010, Sedimentology and sequence stratigraphy of the Hugin Formation, quadrant 15, Norwegian sector, South Viking Graben: Geological Society, London, Petroleum Geology Conference series, **7**, 157–176.

Latour, B., 1986, Visualization and cognition: Thinking with eyes and hands, *in* H. Kuklick, ed., Knowledge and society studies in the sociology of culture past and present: Jai Press 6, 1–40.

Lau, G. T., and S. Ng, 2001, Individual and situational factors influencing negative word-of-mouth behaviour: Canadian Journal of Administrative Sciences/Revue Canadienne des Sciences de l'Administration, **18**, 163–178, doi: 10.1111/j.1936-4490.2001.tb00253.x.

Levin, I. P., 1987, Associative effects of information framing: Bulletin of the Psychonomic Society, **25**, 85–86, doi: 10.3758/BF03330291.

Libarkin, J., and C. Brick, 2002, Research methodologies in science education: Visualization and the geosciences: Journal of Geoscience Education, **50**, 449–455, doi: 10.5408/1089-9995-50.4.449.

Liben, L. S., and S. J. Titus, 2012, The importance of spatial thinking for geoscience education: Insights from the crossroads of geoscience and cognitive science: Geological Society of America Special Papers, **486**, 51–70, doi: 10.1130/9780813724867.

Macrae, E. J., C. E. Bond, Z. K. Shipton, and R. J. Lunn, 2016, Increasing the quality of seismic interpretation: Interpretation, **4**, no. 3, T395–T402, doi: 10.1190/INT-2015-0218.1.

Malkoc, S. A.,, and G. Zauberman, 2006, Deferring versus expediting consumption: The effect of outcome concreteness on sensitivity to time horizon: Journal of Marketing Research, **43**, 618–627, doi: 10.1509/jmkr.43.4.618.

Niccoli, M., 2014, How to evaluate and compare color maps: The Leading Edge, **33**, 910–912, doi: 10.1190/tle33080910.1.

Paton, G. S., and J. Henderson, 2015, Visualization, interpretation, and cognitive cybernetics: Interpretation, **3**, no. 3, SX41–SX48, doi: 10.1190/INT-2014-0283.1.

Patt, A., and R. Zeckhauser, 2000, Action bias and environmental decisions: Journal of Risk and Uncertainty, **21**, 45–72, doi: 10.1023/A:1026517309871.

Polson, D., and A. Curtis, 2010, Dynamics of uncertainty in geological interpretation: Journal of the Geological Society, **167**, 5–10, doi: 10.1144/0016-76492009-055.

Rapp, D. N., and D. H. Uttal, 2006, Understanding and enhancing visualizations: Two models of collaboration between earth science and cognitive science: Special Paper of the Geological Society of America, **413**, 121–127.

Rankey, E. C., and J. C. Mitchell, 2003, That's why it's called interpretation: Impact of horizon uncertainty on seismic attribute analysis: The Leading Edge, **22**, 820–828, doi: 10.1190/1.1614152.

Richards, F. L., N. J. Richardson, C. E. Bond, and M. Cowgill, 2015, Interpretational variability of structural traps: Implications for exploration risk and volume uncertainty: Geological Society, London, special publications 421, 7–27.

Rowbotham, P., P. Kane, and M. Bentley, 2010, Bias in geophysical interpretation — The case for multiple deterministic scenarios: The Leading Edge, **29**, 590–595, doi: 10.1190/1.3422459.

Simons, R. F., B. H. Detenber, J. E. Reiss, and C. W. Shults, 2000, Image motion and context: A between- and within-subjects comparison: Psychophysiology, **37**, 706–710, doi: 10.1111/1469-8986.3750706.

Slothuus, R., and C. H. De Vreese, 2010, Political parties, motivated reasoning, and issue framing effects: Journal of Politics, **72**, 630–645, doi: 10.1017/S002238161000006X.

Sun, Y., S. Li, N. Bonini, and Y. Su, 2012, Graph-framing effects in decision making: Journal of Behavioral Decision Making, **25**, 491–501, doi: 10.1002/bdm.v25.5.

Tversky, A., and D. Kahneman, 1981, The framing of decisions and the psychology of choice: Science, **211**, 453–458, doi: 10.1126/science.7455683.

VanRullen, R., and S. J. Thorpe, 2001, The time course of visual processing: From early perception to decision-making: Journal of Cognitive Neuroscience, **13**, 454–461, doi: 10.1162/08989290152001880.

Wedel, M., and R. Pieters, 2008, A review of eye-tracking research in marketing, *in* N. K. Malhotra, ed., Review of marketing research: Emerald Group Publishing Limited, 123–147.

Welch, B. L., 1947, The generalization of 'Student's' problem when several different population variances are involved: Biometrika, **34**, 28–35.

Welland, M., N. Donnelly, and T. Menneer, 2006, Are we properly using our brains in seismic interpretation?: The Leading Edge, **25**, 142–144, doi: 10.1190/1.2172303.

Zhang, Y., and R. Buda, 1999, Moderating effects of need for cognition on responses to positively versus negatively framed advertising messages: Journal of Advertising, **28**, 1–15, doi: 10.1080/00913367.1999.10673580.

Juan Alcalde received a primary degree (2009) in geology from the University of Salamanca, subsequently completing a master's degree (2011) in geophysics and a doctorate (2014) in earth sciences from the University of Barcelona. He is a geophysicist with interests in reflection seismic processing and interpretation. He has worked in integration and joint interpretation of geophysical data sets, characterization of CO_2 storage sites, and numerical modeling for CCS. More recently, he has conducted parallel projects researching Scotland's potential to develop geothermal and negative emissions technologies, as well as investigating the minimization of biases in seismic interpretation.

Clare Bond received a first-class degree from the University of Leeds followed by a Ph.D. from the University of Edinburgh. She is a structural geologist with interests in interpretational uncertainty and structural modeling. She then worked across a range of sectors and interests, from conservation and policy to technical consulting and academia. She has spent time in industry working for Midland Valley; she initially led their North American and Scandinavian portfolios before managing their client facing knowledge team and has provided geologic input into the development of the structural modeling software, Move. Since 2010, she has held a senior lectureship position at the University of Aberdeen. Her current projects include seismic interpretation, fracture modeling, and the use of LiDAR and photogrammetry to create and interpret virtual outcrop models. She is the chair of the tectonic studies group of the Geological Society of London and she codirects the fold-thrust research group.

Charles Randle received a degree in applied geology from the Camborne School of Mines, and he is finishing a Ph.D. from the University of Aberdeen in partnership with the British Geological Survey. Making use of expertise in geologic modeling and statistical analysis, he is studying the behavior of uncertainty across different geologies and modeling methodologies, along with testing uncertainty quantification methods.