To a High IQ!

Information Presentation Quality:
Assessing Quality of the Communication of Information

by
Larry P. English

INTRODUCTION

This is the third in a series of three articles that describes the important Information Quality (IQ) Characteristics.

As stated in previous articles, the information consumer solely determines what constitutes quality in the information they need to perform their work. Knowledge workers need common IQ characteristics, such as clear, complete and correct definitions, and accurate and complete content. In addition, knowledge workers also need to see that information ***presented*** in a clear, intuitive, and unbiased manner that is appropriate to their functions and the conditions in which they work.

INFORMATION QUALITY[1] CHARACTERISTICS

There are three fundamental *components* or groups of information quality characteristics:

1. Information Definition Quality

Information definition quality is fundamentally the quality of the "information product specifications." Information product specification data includes the data name, definition, valid value set or range, and business rules that represent certain quality requirements.

Information definition is to data (content) what manufacturing product specifications are to the manufactured product.[2] Quality "Information Product Specifications" facilitate the production of consistent quality information.

2. Information Content Quality

Processes that "create" or "update" data produce the "raw materials" of information. Information content quality includes completeness, validity, accuracy, and precision, among others.

3. Information Presentation Quality

When data is retrieved, formatted, aggregated, or combined with other data and presented to knowledge workers, it becomes a finished "information product." Presentation quality characteristics include accessibility, timeliness, presentation intuitiveness, and objectivity, among others.

In this issue we address Information ***Presentation Quality*** characteristics of the communication or presentation of information to knowledge workers.

INFORMATION AS DATA "RAW MATERIAL" AND "FINISHED INFORMATION PRODUCT"

Presentation quality is part of the human-machine interface. Presentation quality characteristics represent the "look and feel" of the finished information product. These characteristics are not just the *prettiness* of information presented, but represent the degree to which the information communicates the message in the data accurately and clearly to the information consumer.

For information to have quality to knowledge workers:

- It must be clearly defined so knowledge workers understand its meaning.
- It must be complete, accurate, and consistent across all data stores.
- It must be accessed and presented in a *timely* basis, and in an *unbiased* way that reveals the truth, so that the knowledge workers can take the right action or make the right decision.

The last set of quality characteristics that knowledge workers require is ***presentation quality*** characteristics, which we discuss in detail below.

INFORMATION PRESENTATION QUALITY CHARACTERISTICS

Quality Characteristics of Information Presentation

Knowledge workers require different content quality characteristics based on their need for that information. Based on my work with dozens of clients, the major information ***presentation*** quality characteristics include:

- **Availability.** Information is accessible when it is needed.
- **Accessibility.** Being able to get the information when needed.
- **Presentation Media Appropriateness.** Being presented in the right technology medium, such as online, hardcopy report, audio, or video.

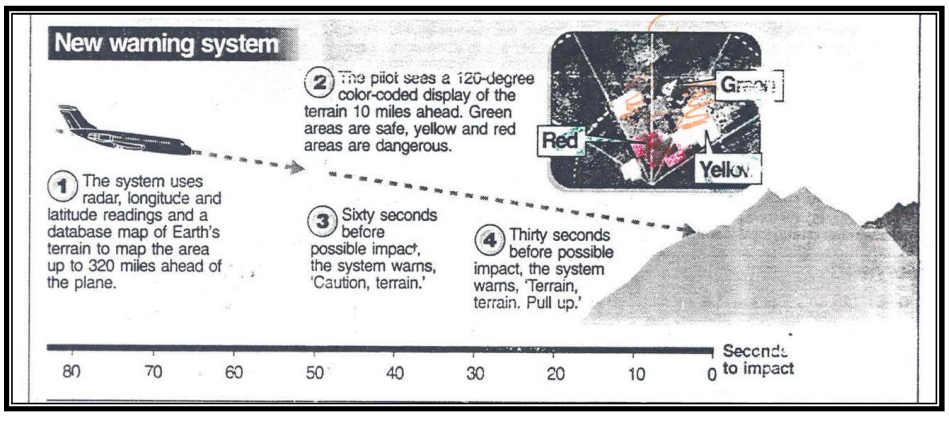

Figure 1. Media Appropriateness in Warning Information Presentation
Warning alerts in the airplane cockpits provide screens with color coding to identify the proximity of the plane to approaching "terrain" (mountains). As a plane approaches the terrain, the color changes from green to yellow to red. If the terrain screen shows red, the plane is about 30 seconds away from impact and warning lights flash as an audio warning message repeatedly calls out, "*Terrain! Pull up!*"

- **Relevancy.** Information is appropriate for the task at hand, i.e., information required to perform a process or make a decision.
- **Presentation Standardization.** Formatted data is presented consistently in a standardized way across different media, such as in computer screens, generated reports, or manually prepared reports.
  - **Structured Values**. Structured attributes like dates, time, telephone numbers, tax id numbers, product codes, and currency amounts should be formatted in a consistent, standard way in any presentation. Numbers and identifiers are easier to remember and use when they are chunked, e.g., standard phone number formats such as +[1] (615) 837-1211.
  - **Structured Documents**. Repeating reports should have a standard format with a style sheet that presents the information in a format that is consistent, easy to read, and easy to understand.

    Documents should use readability-enhancing techniques such as:

    - Information chunking
    - Use of simple words
    - Short sentences with active verbs
    - Bulleted items for lists
    - A readability index of three grade levels below the reading audience
    - Methods such as "Information Mapping" help improve readability of documents. [3]
- **Presentation Clarity.** Information is presented in a way that communicates the truth of the information. Clear labels, footnotes, other explanatory notes, references, or links to definitions and/or documentation that clearly communicate the meaning and any anomalies in the information enhance presentation clarity.

  Changes in data definition or in business rule specification can cause comparing information across time boundaries to be not accurate.

  Figure 2 illustrates one way to represent such changes in the rules for calculating unemployment and how to illustrate it compared to past.

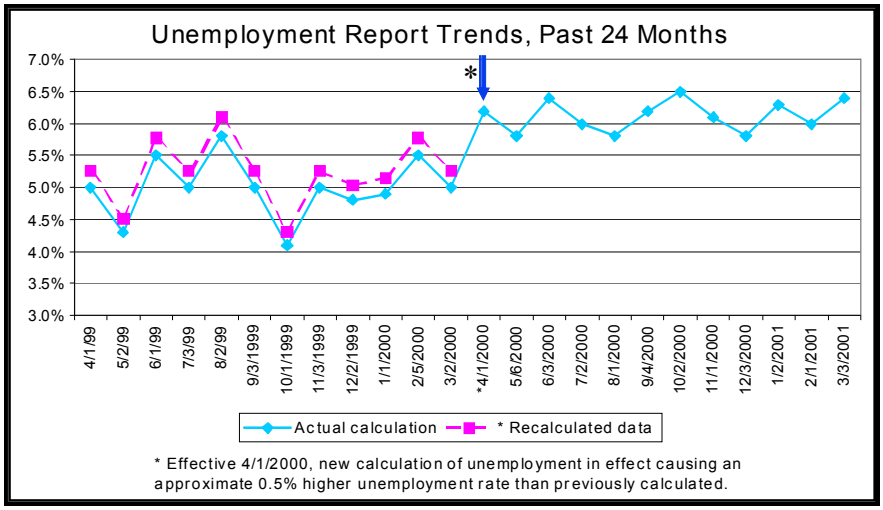

Figure 2. Noting a Calculation Formula Change for Trend Comparison

Figure 3 illustrates ambiguity in the label "Profit" of a mortgage report. "Profit" in this real life example represented the interest profit that would be realized if all mortgages were to be paid through to loan maturity.

Monthly Mortgage Report							
Region	Branch	Num of On-Book Mort	Num of New Mort	Num of Paid Off Mort	Original Principal Amt (in 000s)	Outstanding Principal Amt (in 000s)	Profit (in 000s)
12		7,022	168	167	$1,260,660	$974,947	$5,317,699
	1201	908	12	31	$109,456	$87,565	$470,660
	1202	1,543	37	35	$219,631	$153,741	$900,486
	1205	289	12	18	$102,530	$92,277	$451,133
	1205	1,974	38	22	$348,273	$292,549	$1,480,159
	1206	851	16	19	$99,363	$66,573	$413,349
	1208	1,457	53	42	$381,408	$282,242	$1,601,912
22		6,164	193	182	$1,168,547	$907,537	$4,060,444
	2201	3,308	104	87	$535,155	$438,827	$2,033,589
	2204	2,856	89	95	$633,392	$468,710	$2,026,855
	…	…	…	…	…	…	…

Figure 3. Ambiguous Column Label "Profit"

Monthly Mortgage Adjusted Profitability Report									
Region	Branch	Num of On-Book Mort	Num of New Mort	Num of Paid Off Mort	Original Principal Amt (in 000s)	Outstanding Principal Amt (in 000s)	Max Interest Possible (in 000s)	Net Est Interest * (in 000s)	Monthly Interest Rec'd ** (in 000s)
12		7,022	168	167	$1,260,660	$974,947	$5,317,699	*$2,446,142*	*$9,408*
	1201	908	12	31	$109,456	$87,565	$470,660	*$216,503*	*$833*
	1202	1,543	37	35	$219,631	$153,741	$900,486	*$414,223*	*$1,593*
	1205	289	12	18	$102,530	$92,277	$451,133	*$207,521*	*$798*
	1205	1,974	38	22	$348,273	$292,549	$1,480,159	*$680,873*	*$2,619*
	1206	851	16	19	$99,363	$66,573	$413,349	*$190,141*	*$731*
	1208	1,457	53	42	$381,408	$282,242	$1,601,912	*$736,880*	*$2,834*
22		6,164	193	182	$1,168,547	$907,537	$4,060,444	*$1,867,804*	*$7,184*
	2201	3,308	104	87	$535,155	$438,827	$2,033,589	*$935,451*	*$3,598*
	2204	2,856	89	95	$633,392	$468,710	$2,026,855	*$932,353*	*$3,586*
	…	…	…	…	…	…	…		

Figure 4. Correct Label for Maximum Interest Possible with a New "Estimated Interest" Column

It is far better to correctly label the column and create a column for "estimated net interest profit" based on current loan early payoff experience, as illustrated in Figure 4.

- **Signage Clarity.** Signs and other information-bearing mechanisms like traffic signals should be standardized and made universal across the broadest audience possible.

  Traffic signal lights are now standardized globally with red (stop), yellow (caution), and green (go) meanings. Furthermore, traffic signal lights have standard placements with red on top and green at the bottom for people with color-blindness, so that meaning is consistently associated with the position. This message "redundancy" reduces error.

- **Presentation Objectivity.** Information is presented without bias, enabling the knowledge worker to understand the meaning and significance without misinterpretation.

  Numeric or quantitative data often requires graphical presentation. Objectivity means that the graphical or visual presentation of the information must not distort the truth as evidenced in the data.

For example, Figure 5 shows the difference in bias in a graphical presentation in Zale's 2001 Annual Report. [4]

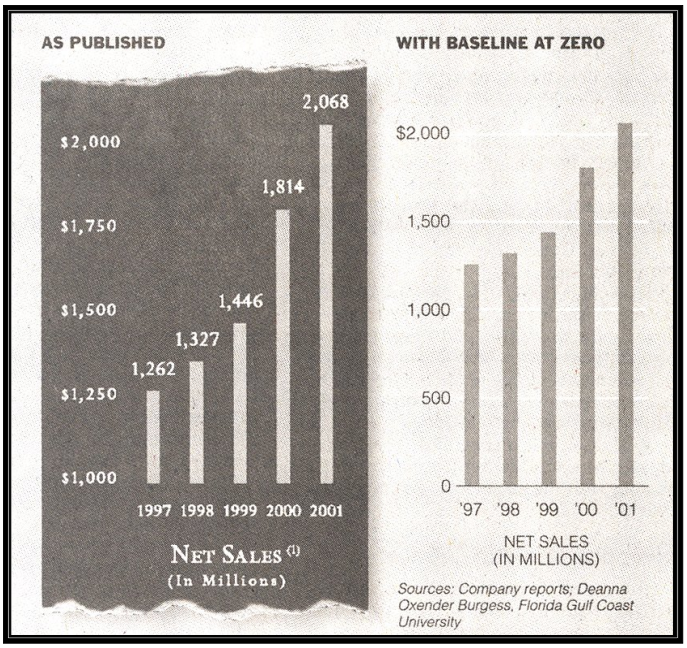

Figure 5. Zale's Net Sales: Report Bias vs. Truthful

The graph of five-year sales is plotted in the report (left graph) with a baseline of $1,000. This baseline biases the data by appearing to show a three-fold growth in sales over five years. The graph on the right presents a truthful picture of sales, plotted with a baseline of zero, accurately represents the "truth" of a sixty-four percent (64%) sales growth.

Edward Tufte measures the degree of distortion in a graphical presentation with what he calls the "Lie Factor," calculated by dividing the size of the effect shown in the graphic by the size of the effect in data. [5]

In the Zales 2001 Annual Report example, the ***size*** of the increase as presented by the ***size*** of the column height "suggests" the 2001 net sales is 3.8 times the 1997 net sales. This distortion is created by setting the baseline at the $1,000 million mark — not an origin of zero.

The graph on the right, on the other hand, accurately uses the size of the bars to depict the net sales data accurately. The five-year net sales increase was a still healthy, sixty-four percent (64%).

To calculate the Lie Factor, first calculate the sales increase represented in the data itself, then calculate the sales increase represented by this graphic representation.

In our example, we measure the height of the first column representing $1,262. In the original the height is 20.5 mm. Measure the size of the last column representing $2,068. The original height is 78.5 mm.

Increase represented by data = ((2,068 - 1,262) / 1,262) X 100 = **64%**

Increase represented by graphic = ((78.5 – 20.5) / 20.5) X 100 = **283%**

The Lie Factor is then calculated by dividing the size of the effect shown in the graphic by the size of the effect shown in the data:

Lie Factor = 283% / 64% = **4.4**

The graphic makes the end result appear to be 4.4 times larger than the actual increase shown in the data itself.

A Lie Factor greater than 1.05 or less than 0.95 is a critical problem because humans assimilate visual information much faster than descriptive information. Such representations may therefore mislead people into making wrong decisions.

- **Presentation Utility.** Information is presented in a way that is intuitive and appropriate for the task at hand. The presentation of information will vary by the individual uses for which it is required. Some uses require concise presentation, while others require a complete, detailed presentation, and yet others require graphics, color-coding, or other highlighting techniques.

MEASURING INFORMATION PRESENTATION QUALITY

Different quality characteristics require different measurement techniques. Some presentation quality characteristics are measured subjectively by knowledge workers who rate their satisfaction based upon the ease of use and speed with which they assimilate it into their work.

You can, however, design experiments to statistically measure presentation quality. For example, you can design measurements of how long it takes to use information presented in different formats.

Take a time demonstration test of locating data in a paragraph structured document versus locating information in an information-mapped document at: http://www.infomap.com/index.cfm/Expertise/Brochures and click on the box labeled “IMAP Demo” at the bottom of the lower left of the page.

Tests can be developed to measure the percent of “wrong” actions or “wrong” decisions (when the correct decision is knowable from available information) made by groups of knowledge workers using the different information presentation forms.

GUIDELINES FOR INFORMATION PRESENTATION QUALITY OF GRAPHICAL INFORMATION

Edward Tufte provides us with some guidelines for presenting quantitative information visually or graphically.

Guides for visual quality[6]

- Choose a proper format and design
- Use words, numbers, and drawings together
- Reflect a balance, a proportion, a sense of relevant scale
- Display an accessible complexity of detail
- Often have a narrative quality, a story to tell about the data
- Draw in a professional manner, with the technical details of production work done with care
- Avoid content-free decoration, including chartjunk, i.e., graphical designs or artistic decoration that is "non-data ink" and which does not communicate the message in the data

Guides for Knowledge-Worker Friendly Graphics[7]

The table below compares friendly and unfriendly characteristics of graphics.

Friendly	Unfriendly
Words are spelled out, mysterious and elaborate encoding avoided	Abbreviations abound, requiring the viewer to sort through text to decode abbreviations
Words run from left to right, the usual direction for reading occidental languages	Words run vertically, particularly along the Y-axis; words run in several different directions
Little messages help explain data	Graphic is cryptic, requires repeated references to scattered text
Elaborately encoded shadings, cross-hatching, and colors are avoided; instead, labels are placed on the graphic itself; no legend is required	Obscure codings require going back and forth between legend and graphic
Graphic attracts viewer, provokes curiosity	Graphic repels the viewer, lots of "chartjunk"
Colors, if used, are chosen so that the color-deficient and color-blind (5 to 10 percent of viewers) can make sense of the graphic (blue can be distinguished from other colors by most color-deficient people)	Design is insensitive to color-deficient viewers; red and green are used for essential contrasts
Type is clear, precise, modest; lettering may be done by hand	Type is clotted, overbearing
Type is upper-and-lower case, with serifs	Type is all capitals, sans serif

Principles of Graphical Excellence [8]

- Graphical excellence is the well-designed presentation of relevant data — a matter of *substance*, of *statistics*, and of *design.*
- Graphical excellence consists of complex ideas communicated with clarity, precision and efficiency.
- Graphical excellence is that which gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space.
- Graphical excellence is nearly always multivariate.
- Graphical excellence requires telling the truth about the data.

Principles of Graphical Integrity [9]

- The representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities represented.
- Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data.
- Show data variation, not design variation.
- In time-series displays of money, deflated and standardized units of monetary measurement are nearly always better than nominal units.
- The number of information-carrying (variable) dimensions depicted should not exceed the number of dimensions in the data.
- Graphics must not quote data out of context.

CONCLUSION

The presentation of information to knowledge workers represents the final "manufacturing" process of information by assembling the information in a form and way that communicates the information clearly to the customer. Information presentation cannot improve the quality of the data "content." Information presentation cannot improve the "definition" or "meaning" of data. It can, however, bring the truth in the data to light.

Information presentation must "tell the truth" in the data without bias, and put information in a context that is clear and understandable to the knowledge worker. Information presentation must also call attention to any anomalies in the data, such as degree of missing and inaccurate data, so that knowledge workers can apply their knowledge to take the right action or make the right decision.

For more about information content quality characteristics and how to measure them please see Chapter 6, "Information Quality Assessment" in *Improving Data Warehouse and Business Information Quality*.[10]

To a High IQ!!! Let me hear about your experiences at larry.english@infoimpact.com

About the Author

Larry P. English, Co-Founder of the IAIDQ, is President and Principal of INFORMATION IMPACT International, Inc., Brentwood, TN., and author of *Improving Data Warehouse and Business Information Quality*. He provides consulting and training to help Information Professionals increase their value to the enterprise. Mr. English provides certification in his TIQM methodology. For details, email TIQMCert@infoimpact.com.

End Notes

[1] The terms "Information Quality" and "Information Quality Management" are used as synonyms with "Data Quality" and "Data Quality Management."

[2] Larry P. English, *Improving Data Warehouse and Business Information Quality*, (*IDW&BIQ*) NY: John Wiley, 1999, p. 84.

[3] Information Mapping, Inc., Home page last accessed 7 Oct, 2006 at: http://www.infomap.com/index.cfm/

[4] Deanna Oxender Burgess, "Exaggerating the Positive," *The New York Times*, September 15, 2002.

[5] Edward Tufte, *The Visual Display of Quantitative Information, 2nd Ed.*(*Visual Display*), Cheshire, Conn: Graphics Press, 2001, p. 57.

[6] Tufte, *Visual Display*, p. 177.

[7] Tufte, *Visual Display*, p. 183.

[8] Tufte, *Visual Display*, p. 51.

[9] Tufte, *Visual Display*, p. 77.

[10] English, *IDW&BIQ*, pp. 137-197.