What Have We Learned from 30 Years of Posters? An Assessment Instrument for Posters

C.A. Watkins[1], D.M. VanLeeuwen[2], and J.G. Mexal[3]
New Mexico State University
Las Cruces, NM 88003

Abstract

Posters have been used for over 30 years at scientific meetings, and their use is growing. While they have become a major form of communication at scientific meetings, science fairs and classrooms, few instruments exist to evaluate posters and little has been done to assess the reliability of existing instruments. The objective of this work was to evaluate an instrument and describe the components of a well-designed poster. The instrument separates the components of a poster into the Visual and Content components. Using this instrument, the instructor could use as few as three evaluators to construct a reliable evaluation. Only six evaluators would be needed for a high degree of reliability.

Introduction

The use of posters at scientific meetings first began in 1974 at the Biochemistry-Biophysics conference where 500 of the 2200 presentations were poster presentations (Maugh, 1974). Since then, the popularity of posters has grown steadily. For example, the use of posters at the American Society of Mammalogists conference has grown 1.6% per year to about 40% of presentations since their introduction in 1979 (Genoways and Freeman, 2001). Other national conferences have an even higher proportion of poster presentations. The American Society of Agronomy averages 55% (K.R. Schlesinger, personal communication, 9/10/04), while the American Society for Horticulture Science had 60% posters in 2003 (Anon., 2003). The 2005 annual conference of the Association of College and Research Libraries had 68% poster presentations (Anon., 2005), while the Society for Neuroscience had 88% poster presentations at the 2004 meeting (Anon., 2004). Regional meetings which tend to be smaller, may have a greater percentage of oral presentations, but the use of posters is increasing in those meetings as well.

The popularity of posters is a function of the growth of the societies, and with the need for concurrent sessions, it is not uncommon for some conferences to have over 1,000 presentations. The preference of attendees to view posters (Davis et al., 1992) and the opportunity to reuse posters at other meetings or to publicize research accomplishments on the home campus (Lyons et al., 1985) are reasons for the increased popularity. Posters are viewed not only as important communication tools for scientists, but also as valuable learning tools for both undergraduate and graduate learning.

Undergraduate Study

Posters have been identified as a useful undergraduate learning tool for many of the science disciplines. Fields such as geology (Kemp and Clark, 1992), nursing (Collins, 1992), organic chemistry (Huddle, 2000), and physiology (Fourtner et al., 2002) have used the poster presentation format to teach course content. Students learn both the skill to prepare and present a poster as well as learn the concepts of the research presented in the particular article (Denzine, 1999). As student learning and comprehension improve, it is likely that course material retention increases as well (Smith, 2002). Increased preparation from earlier courses would increase success in advanced courses.

Graduate Study

Posters also have become an integral part of graduate student education. Students that complete a research thesis often are encouraged to present their findings at scientific meetings. The sheer growth of these meetings dictates that student presentations likely will be accepted in poster format rather than as an oral presentation. Both formats can help the student select the most important or critical research findings so the results can be presented in an abbreviated fashion (either 12 minutes for an oral presentation or a readable text format for posters).

Young Faculty

Finally, posters can be an important organizational tool, helping a young faculty member structure the research findings into a concise format. Young scientists must realize that posters are intended to be 'expanded abstracts' rather than 'mini-journal articles' (Gosling, 1999). Nevertheless, the ability needed to prepare a poster can facilitate the preparation of the manuscript.

While it is obvious that posters are becoming an ever increasing form of scientific communication, there is little information on the quantitative description of factors that constitute a 'good' poster.

[1]Associate Professor, Library
[2]Professor, Department of Agricultural and Extension Education
[3]Professor, Department of Agronomy and Horticulture

Table 1. Poster Evaluation Questionnaire (Modified from Hess and Brooks, 1998)

60-Second Poster Evaluation

Poster Title: ____

1. Overall Appearance

_____ 0 Cluttered or sloppy appearance. Gives the impression of a solid mass of text and graphics, or pieces are scattered and disconnected. Little white space.
_____ 1 Pleasant to look at. Pleasing use of colors, text and graphics.
_____ 2 Very pleasing to look at. Particularly nice colors and graphics.

2. White Space

_____ 0 Very little. Gives the impression of a solid mass of text and graphics.
_____ 1 OK. Sections of the poster are separated from one another.
_____ 2 Lots. Plenty of room to rest the eyes. Lots of separation.

3. Text/Graphics Balance

_____ 0 Too much text. The poster gives an overwhelming impression of text only.
_____ 1 OK.
_____ 2 Balanced. Text and graphics are evenly dispersed in the poster, enough text to explain the graphics.

4. Text Size

_____ 0 Too small to view comfortably from a distance of 1-1.5 meters.
_____ 1 Main text OK, but text in figures too small.
_____ 2 Easy to read from 1-1.5 meters.

5. Organization and Flow

_____ 0 Cannot figure out how to move through poster.
_____ 1 Implicit. Headings (I, M, R and D) or other device implies organization and flow.
_____ 2 Explicit numbering, column bars, row bars, etc.

6. Author Identification

_____ 0 None.
_____ 1 Partial. Not enough information to contact author without further research. Small font size
_____ 2 Complete. Enough information to contact author without further research. Good font size.

7. Research Objective

_____ 0 Can't find.
_____ 1 Present, but not explicit. Buried.
_____ 2 Explicit. This includes headings of "Objectives", "Aims", "Goals", etc.

8. Main Points

_____ 0 Can't find.
_____ 1 Present, but not obvious. May be imbedded in monolithic blocks of text.
_____ 2 Explicitly labeled (e.g., "Main Points", "Conclusions", "Results").

9. Clarity of Information

_____ 0 Limited jargon was used so persons outside the field can understand the content
_____ 1 Content was fully explained
_____ 2 Concepts were defined when definitions were needed.

TOTAL SCORE ________ **pts**

Instruments for judging posters are available, but there has been no validity or reliability evaluation of these instruments. The objectives of this manuscript are to evaluate an instrument used for judging posters and using the evaluator's scores, describe components of a well-designed poster.

Materials and Methods

Fifteen posters (34 X 44 inches), prepared by faculty as part of their research and students as part of a research orientation course (Sammis and Mexal, 1996) were evaluated by 11 professionals using an instrument modified from Hess and Brooks (1998) (Table 1). Originally items had inconsistent response scales. Modifications were made to obtain a three point scale on all items to accommodate forming summated scores the researchers hoped would be reliable. Posters were intentionally selected to provide a range of color, text, and graphics. The evaluators were librarians specifically interested in poster presentations attending the LOEX (Library Orientation Exchange) of the West Conference, 2002. The evaluators were given a brief presentation of the components of an effective poster, and then viewed each poster for about 60 seconds as they graded the nine questions.

This study was designed to provide insight into the performance of this rating instrument. Consequently, posters were selected representing a range of quality. In addition, Generalizability Theory (GT), which involves estimation of variance components in random models, requires a reasonably large number of raters and posters.

As a first step in evaluating the instrument, responses were subjected to a confirmatory principal component factor analysis (Johnson, 1998) in order to confirm, or possibly refute, researchers' prior beliefs about instrument subscales. A scree plot was used to determine the number of factors, and the initial factor pattern was subjected to a varimax rotation to produce interpretable factors. For each confirmed subscale, poster measurement reliability based on summing across items and raters was assessed using Generalizability Theory (VanLeeuwen, 1997; Shavelson and Webb, 1991). Variance components were estimated using method of moments estimators based on type III sums of squares (SAS Institute, 1999). All analyses were executed using SAS version 8.2.

Results and Discussion

Of the nine questions (Table 1), Q-1 to 4 related to the visual components of the poster, while Q-5, 7 to 8 related to the content factor of the poster. A scree plot suggested the presence of either two or three factors. When two components were retained and a varimax rotation applied to the factors, Q-1, 2, 3, and 4 loaded highly on the first component (Visual). All four question's loadings were greater than 0.70, whereas first factor loadings were smaller than 0.25 for the other questions. Q-5, 7, and 8 loaded highly on the second factor (Content). Again, all three factor loadings were greater than 0.70. These factors were consistent with a subscale assessing a poster's Visual aspects and a second subscale assessing Content, which were consistent with the researcher's prior conjecture that items 1 to 5 might form one scale and items 5,7,and 8 another. The factor analysis places item 5 only in the second scale. Q-6 did not load highly on either factor. The rotated factor pattern when retaining three factors was similar. Q-6 had a large loading only on the third factor. Therefore, Q-6 was removed from the analysis. Q-9 also was removed from the confirmatory factor analysis because more than half of the respondents did not rate this item. The final confirmatory factor analysis categorized the items into two scales: Visual and Content (Table 2). The Visual scale consisted of Q-1 to 4; the Content, Q-5, 7, 8; and a third scale combined all seven items into the Overall poster rating.

Table 3 summarizes estimates of the relative decision variance and G coefficient (reliability) estimates (VanLeeuwen, 1997). The G coefficient is analogous to a traditional reliability coefficient assessing the reliability of relative decisions. The error in relative decisions is affected by interactions involving posters as well as measurement error. Conceptually, the reliability of a set of poster

Table 2. Analysis of Variance Component Estimates for Visual, Content, and Overall Scores

Source of Variance	Visual (Q1-Q4)		Content (Q5, Q7, Q8)		Overall (Q1-Q5, Q7, Q8)	
	Est.	%	Est.	%	Est.	%
Poster	0.2625	37.96	0.1727	31.09	0.1026	15.60
Rater	0.0586	8.47	0.0625	11.25	0.0593	9.02
Item	-0.0037	0	0.0089	1.60	0.0309	4.70
Poster * Rater	0.1014	14.66	0.0514	9.25	0.0580	8.82
Poster * Item	0.0650	9.40	0.0617	11.11	0.1829	27.82
Rater * Item	0.0251	3.63	0.0108	1.94	0.0191	2.90
PRI, error	0.1790	25.88	0.1874	33.74	0.2047	31.13

Table 3. Reliabilities and Standard Deviation for Visual, Content, and Overall Scores, Given Varying Numbers of Raters

Number of Raters	Visual (Q1-Q4)		Content (Q5, Q7, Q8)		Overall (Q1-Q5, Q7, Q8)	
	σ^2rel	Reliability	σ^2rel	Reliability	σ^2rel	Reliability
11	.0295	0.90	.0309	0.85	.0341	0.75
6	.0406	0.87	.0395	0.81	.0407	0.72
5	.0455	0.85	.0433	0.80	.0436	0.70
4	.0528	0.83	.0490	0.78	.0479	0.68
3	.0650	0.80	.0585	0.75	.0552	0.65

measurements is the proportion of the variability in the measurements that is due to posters. That is, it is the ratio of the poster variance to the sum of the poster variance and the error variance. Since a poster measurement is based on averaging across items and raters, the more items and/or raters, the more reliable the measurement. More specifically, the relative decision error is the sum of (the poster by rater interaction variance component divided by the number of raters), (the poster by item interaction variance component divided by the number of items), and (the error variance component divided by the product of the number of items and the number of raters).

For the Visual scale, the poster variance component is large relative to the other components and accounts for nearly 38% of the total variance (Table 2). The error also accounts for a high percentage of the variance for all three scales. Only in the Overall scale did the Poster*Item variance component exceed the poster component and approach the magnitude of the error component.

A Visual rating based on the four questions and 11 raters was estimated to have a reliability of 0.90 (Table 3). A Visual rating based on only three raters was estimated to have a reliability of 0.80, which is still acceptable. Although reliabilities for the Content measurement are reasonably high (Table 3), they were slightly lower than the reliabilities for the Visual measurement. Relative to the other variance components involved in relative decisions, the poster variance component was slightly smaller for this scale than for the Visual scale. Nonetheless, depending on the application, reliabilities as low as 0.70 may be acceptable; based on the estimates and data from this study, measurements using only three raters may provide a sufficiently reliable ranking for the Visual and Content subscales.

Reliabilities for the Overall scale are lower than those for either of the subscales. Looking at the variance estimates and percentages provides insight into the combined scores' reduced reliability. Compared to either of the subscales, the Overall scale's poster variance component accounted for a smaller percentage of the total variance in a single response. However, the poster by item interaction variance component accounts for substantially more of the variance. This occurs because the Visual and Content components do not necessarily go up and down together. When we pool all the items together, this leads to some reduction in the poster variability as well as an increase in the poster by item variance component. That is, posters that are high on Visual and low on Content (or vice versa) tend to have mediocre Overall scores that reduce the poster variability but contribute substantially to the higher poster by item interaction. Even though averaging across seven rather than three or four questions balances out the higher interaction variance, the reduced poster variance brings the reliability of the measurements down.

Table 4 summarizes the ratings for the posters in this study. Posters are listed in increasing order of the Overall scale score. Note that the order would be somewhat different if the list were ordered by either of the subscales. For example poster 7 had the highest Visual rating of 1.73 while its Content rating was only 0.91, substantially less than the highest Content score of 1.82. The combined score fell somewhere in the middle.

GT departs somewhat from classical reliability in that it provides the basis for standard error estimates that can be used to provide guidance as to meaningful difference magnitudes. For example, consider the Visual rating based on 11 raters. The variance (Table 3) was 0.0295. Consequently, the variance of the difference between two posters' Visual scores is estimated to be 0.059 (=2x0.0295) and the standard error for the difference is obtained by taking the square root of 0.059. Because this square root, and therefore the standard error of the difference, is 0.243, a difference in poster Visual scores of only 0.243 may be due to chance error alone. Furthermore, a more conservative approach would be to require that the difference between posters exceed 2x0.243 =0.486 to consider the poster Visual difference to be larger than might reasonably be attributed to error variability alone (VanLeeuwen et al., 1999). Consequently, only posters having Visual scores less than 1.244 might be considered to have a lower Visual ranking than poster 7.

Table 4. Mean Scores for Visual, Content and Overall Ranking Ordered on Overall Score

Poster	Visual Score	Content Score	Overall Score
5	0.09848	0.54545	0.29545
15	0.70455	0.68182	0.70606
11	0.06818	1.57576	0.71429
6	0.63636	1.00000	0.79221
13	0.72727	1.30303	0.97403
12	0.47727	1.78788	1.03896
1	0.79545	1.69697	1.18182
3	1.63636	0.66667	1.22078
4	0.84091	1.81818	1.25974
7	1.72727	0.90909	1.37662
10	1.18182	1.72727	1.41558
9	1.29545	1.66667	1.45022
14	1.56818	1.39394	1.49351
2	1.40909	1.72727	1.54545
8	1.50000	1.66667	1.57143

Components of well-designed posters

When planning a poster for a conference, the conference often provides instructions on preparing a poster. These tend to focus on the mechanics of a poster, such as: time (including length of session), poster size, abstract, and poster number (Anon., 2001). They may also suggest font size for various components. However, a successful poster requires more. Making a poster should involve two goals: a) creating an interesting visual appearance that will encourage colleagues to read the poster, and b) providing straightforward content that is easily comprehended. In order to achieve these goals the design of the poster takes on greater importance.

To compare components of low and high rated posters, two posters used in the analysis, illustrated in Figure 1, are summarized in Table 5. Poster 2 (Table 4) scored high in both categories and therefore, overall, and poster 11 scored low in Visual and consequently Overall, but had a relatively high Content score.

Table 5. Summary characteristics of posters with high and low overall scores (see also Figure 1)

Poster 2 (High Score)	Poster 11 (Low Score)
Pleasing, soft color background	Plain white background
Lots of space, 466 words (3 columns)	Text (1,647 words) and figure rich (4 columns)
Active, short title (7 words), lower case letters	Long (23 words), all capital letters
Prominent author identification	Buried, small font author identification
Large, sans serif font (32 pt)	Small font (18 pt)
Clear, serif headings	Clear but small headings
Objectives clearly stated	Objectives buried
Prominent, highlighted conclusion	Smothered conclusion
Large, easy to read graphics	Easy to read graphics, but little explanation
Graphics (5) linked to text	Graphics (9) linked to text (too many?)
Large, relevant photo	Good photos, but small
Prominent, highlighted conclusion	Smothered conclusion

Visual component

The most significant aspect relating to visual appearance is limiting the amount of text on the poster. Poster 2 had 466 words, whereas Poster 11 had 1,647 words. Poster designers should ask themselves is there a way to visually illustrate this text. Gosling (1999) suggests that a poster designer consider "expanding the abstract rather than condensing the paper." Blocks of text add to the density of a poster, while margins and line spacing add to the openness or white space. Text alignment should be right-ragged and line spacing should be 1.5 or as open as possible. A large font size (> 24 pt) is a good self-limiting tool. Poster session attendees will be viewing the poster from possibly several meters. Therefore the font of the poster needs to be large and easy to read. A 24-point (Helvetica) font for the text and a 96-point font for the title are suggested. A sans serif font for the text of the poster and a serif font for the title and headings adds subtle visual contrast (Williams, 1994).

Figures, photos, or diagrams should be large enough to be easily read from 2 m, and should be linked to the text. Two of the top three visually rated posters (Posters 7, 14) had pictures which included people working. This is an often-overlooked aspect of data collection. However, as presentations shift from oral to poster, action pictures may attract more attention.

Color is probably the most striking feature of a good poster. However, the color combinations should not be overpowering, but chosen carefully for a well designed poster. Nicol and Pexman (2003) advise choosing two or no more than three colors and keeping them consistent throughout the poster.

Content component

The title is one of the first 'hooks' to attract readers. It should be attention grabbing, and convey the underlying message of the results. Avoid empty words, such as: influence, affect, study. Avoid using all capital letters, which can be difficult to read. If the title is short, an exotic, script font can be used effectively. Font size should be at least 72 pt (~3/4"). Author identification and affiliation should be only slightly smaller than the title, and easily discernable (Figure 1A).

Main headings should be clearly labeled in a font larger than the text. Complete sentences should be minimized and bulleted text used wherever possible. It is difficult to 'defy gravity' in organization and flow of major sections. However, important headings, such as 'Objectives' and 'Conclusions' should neither be buried in text (Figure 1B) nor placed in the bottom

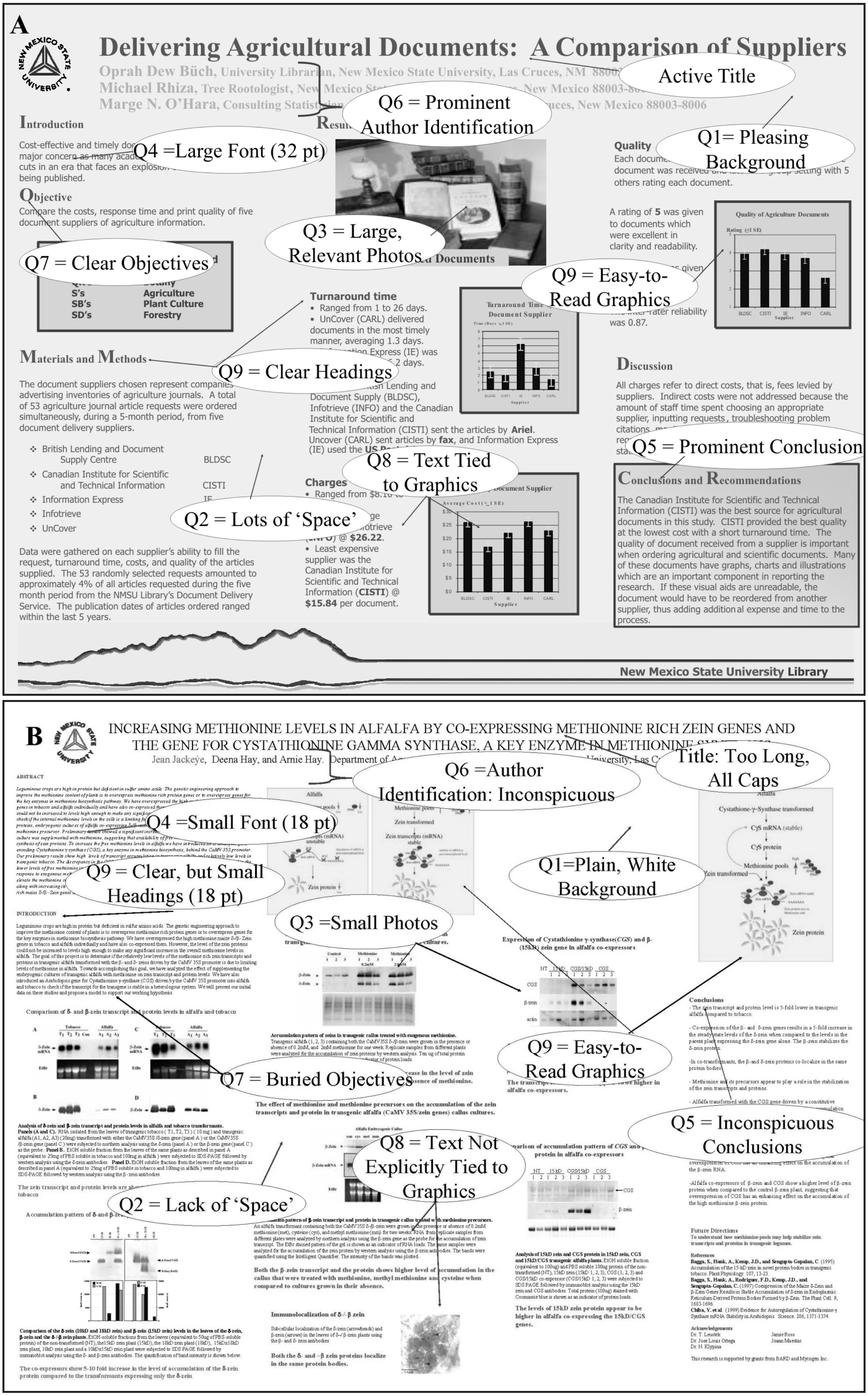

Figure 1. A comparison of two posters.

corners of the poster. Research conclusions should be in a prominent position, if possible or placed in a distinguishing text box to readily set it off from the body of the text.

Photos or diagrams are components to illustrate key points. Large, complicated tables of data can be acceptable for manuscripts, where the reader can take time to comprehend major findings. This is not possible in what may be a crowded poster display. Consequently, the author should design graphics that quickly illustrate a few key findings. It is not necessary to present all the data collected as part of an experiment. If data are presented in tables, the tables should be simple (two to three columns by four to five rows). Likewise, figures should be easily understood. Bar charts or line graphs with few elements will enhance understanding. Furthermore, diagrams or graphics should be equally easy to understand (Figure 1B).

Printing the poster

The simplest approach to printing a poster is to develop the presentation in large format from the outset. In Powerpoint®, the author should go to 'File', 'Page Setup', and select 'Custom page (size 44 inches wide x 34 inches long) with Landscape Orientation. Activating 'Guides' under the 'View' menu will help in text and figure alignment. Photos, graphs or figures can be imported or constructed directly within the poster 'slide'. Creating separate text boxes for major headings will aid in formatting and placement. The draft poster can be printed (after scaling to fit paper) on 8.5 x 11- inch paper for proof-reading or design changes. This is also a good time to check for readability. If the text can be easily read on paper, it will likely be easy to read in the poster. Many university campuses and commercial copy companies (e.g. Kinko's) can print for a reasonable fee. However, the authors should budget to print twice, as often the larger format uncovers errors or additions that are needed before attending the conference.

Summary and Recommendations

This instrument is one of few used to evaluate posters that has been subjected to statistical reliability analysis. The analysis indicated several points. The first was that Q-6 did not fit into either the Visual or Content component but may have value as a check-list item because posters should include author identification. The second was that Q-9 was a poor question that did not fit easily into either subscale. The third point was the questions should be organized into separate Visual and Content categories. Finally, using this evaluation form, instructors could use as few as three evaluators to achieve sufficient reliability, with as few as six to achieve a desirable reliability for the Visual and Content subscales.

Many teachers use posters as learning tools for undergraduate students. This evaluation instrument would allow the instructor to evaluate separately, or exclusively, the Content subscale vis-à-vis the Visual subscale. Additionally, rather than simply averaging all of the items, when obtaining an Overall or combined score, instructors could weight the Content subscale more heavily than the Visual. Regardless, this instrument provides some degree of reliability to judging or grading posters.

Literature Cited

Anon. 2001. Presentation guidelines for ASHS-2001. ASHS Newsletter 17(6):12.

Anon. 2003. American Society for Horticultural Science Centennial Conference Issue-Program and

Abstracts. HortScience 38(5):650-867. and Anon. 2004. Neuroscience 2004. final program. 34th Annual Meeting Society for Neuroscience, San Diego, CA. Oct. 23-27. http://web.sfn.org/am2004/prog2004.pdf

Anon. 2005. Association of College and Research Libraries. 12th National Conference. http://www.ala.org/ala/acrl/acrlevents/12thnatconf/acrlprogram/program.htm

Collins, B.A. 1992. The poster session as a learning experience for master's students. Nurse Educator 17(3):23, 27, 31.

Davis, M., K.J. Davis, and D.C. Wolf. 1992. Effective communication with poster displays. Jour. of Natural Resources and Life Sciences Education 21(2):156-160.

Denzine, G.M. 1999. An example of innovative teaching: Preparing graduate students for poster presentations. Jour. of College Student Development 40(1):91-93.

Fourtner, C.R., M. Bisson, and C.A. Loretz. 2002. Using posters in case studies: The scientific poster as a teaching tool. (http://ublib.buffalo.edu/libraries/projects/cases/posters.html.) (2/11/02).

Genoways, H.H. and P.W. Freeman. 2001. Evolution of a scientific meeting: Eighty annual meetings of the American Society of Mammalogists, 1919-2000. Jour. of Mammalogy 82(2): 582-603.

Gosling, P.J. 1999. Scientist's guide to poster presentations. New York: Kluwer Academic /Plenum Pub.

Hess, G.R. and E.N. Brooks. 1998. The class poster conference as a teaching tool. Jour. of Natural Resources and Life Science Education 27:155-158.

Huddle, P.A. 2000. A poster session in organic chemistry that markedly enhanced student learning. Jour. of Chemical Education 77(9)1154-1157.

Johnson, D.E. 1998. Applied multivariate methods for data analysis. New York, Duxbury Press.

Kemp, K.M., and J.A. Clark. 1992. Teaching geology using poster assignments. Jour. of Geological Education 40:398-403.

Lyons, R.E., T.A. Fretz, and R.T. Johnson. 1985. Poster presentations: An update. Hortscience 20(1): 15-16.

Maugh II, T.H. 1974. Poster sessions: A new look at scientific meetings. Science 184:1361.

Nicol, A.A.M., and P.M. Pexman. 2003. Displaying your findings: A practical guide for creating figures, posters, and presentations. Washington, DC: American Psychological Association.

Sammis, T.W. and J.G. Mexal. 1996. Teaching the components of agricultural research. Jour. of Natural Resources and Life Sciences Education 25(1):58-61.

SAS Institute Inc. 1999. SAS/STAT® software: Changes and enhancement through release 8.2. Cary, NC: SAS Institute Inc.

Shavelson, R.J. and N.M. Webb. 1991. Generalizability theory: A primer. Newbury Park, CA: Sage Publ., Inc.

Smith, C.N. 2002. Using the cell signaling literature to teach molecular biology to undergraduates. Biochemistry and Molecular Biology Education 30(6):380-383.

VanLeeuwen, D.M. 1997. Assessing reliability of measurements with generalizability theory: An application to inter-rater reliability. Jour. of Agr. Education 38(3):36-42.

VanLeeuwen, D.M., T.J. Dormody, and B.S. Seevers. 1999. Assessing the reliability of student evaluations of teaching (sets) with generalizability theory. Jour. of Agr. Education 40(4):1-9.

Williams, R. 1994. The non-designer's design book: Design and typographic principles for the visual novice. Berkeley, CA: Peachpit Press.