Development of a biophotonics concept inventory for program evaluation

Jodie A. Galosy[a], Marco Molinaro[a], Michelle McCombs[a], Ana Corbacho[a]
[a]Center for Biophotonics Science and Technology, University of California, Davis

ABSTRACT

The Center for Biophotonics Science and Technology has developed an evaluation tool to assess the impact of its educational programs on participants' understanding of basic concepts underlying biophotonics science. The Biophotonics Concepts Inventory (BPC) includes fourteen items; some adopted from other concept tests as well as several original items developed by CBST scientists and educators. Scientists also contributed to instrument development by completing a pilot version of the BPC during the CBST annual retreat and rating each item for relevance and importance to the field of biophotonics. The final items were selected based on item feedback and comparisons between mean item scores for scientists, undergraduates, and high school students who completed the draft version of the BPC. Items primarily focus on the behavior of light and light-matter interactions. The instrument has been used as a pre-test and post-test in programs for undergraduates, K-12 teachers, high school and middle school students. To date, there has been a significant increase in BPC scores from pre to post conditions across CBST programs and courses. We will discuss BPC development, response patterns, and pre/post group comparisons. Specifically, we will address how typical misconceptions about light and light-matter interactions were used to design items, the rationale for incorporating visual representations into many items, the methods used for investigating instrument quality, and implications for making claims about the effectiveness of CBST biophotonics education programs.

Keywords: Biophotonics, Education, Evaluation, Concept Inventory, Grades 7-14

1. INTRODUCTION

The Center for Biophotonics Science and Technology (CBST)—a Science and Technology Center (STC) funded by the National Science Foundation (NSF)—advances research, development, and application of new optical/photonic tools and technology in medicine and the life sciences. The core mission of CBST Education and Outreach is to engage diverse audiences with the interdisciplinary science of biophotonics-the study of life with light. CBST Education programs include research academies for high school students, professional development workshops for teachers in grades 6-14, and research experiences for undergraduate interns. One of the principal challenges in evaluating the effectiveness of these programs is measuring impact on participants understanding of biophotonics. This paper documents the process, to date, of developing a tool –the Biophotonics Concept Inventory (BPC)—to assess the impact of program activities on basic concepts key to biophotonics.

CBST scientists have been essential partners in developing the BPC; something that would not have been possible without the interaction of scientists and educators afforded by NSF requirements that STCs have a substantive educational component. CBST educators and scientists collaborated in identifying major basic concepts underlying biophotonics, consulting on item development, and validating the BPC instrument's pilot version. The resulting instrument, although still in development, has proven, thus far, to be a relatively trustworthy and useful tool for evaluating program impact on participants' concept understanding.

The purpose of this paper is to describe BPC development, validation, and results to date as well as discuss future BPC improvement and implications for developing similar tools in other optics education areas.

2. METHODOLOGY

Prior to BPC development, evaluation of program impact on participants' understanding of biophotonics was accomplished with a set of twenty multiple-choice questions. The questions were written by the CBST Education Team

Optics Education and Outreach, edited by G. Groot Gregory, Proc. of SPIE Vol. 7783,
778308 · © 2010 SPIE · CCC code: 0277-786X/10/$18 · doi: 10.1117/12.862355

(CBST Ed), in consultation with an external evaluator, to reflect the information provided in courses and workshops. Although participants' scores generally improved, there were some concerns about instrument quality. These included the extent to which the items: (1) Captured conceptual understanding rather than memorization; (2) Ranged in difficulty; (3) Were relevant and important to biophotonics; and (4) Could be understood by a range of diverse learners.

In order to gather data about these concerns, CBST Ed administered the items at the 2007 CBST summer retreat—an annual gathering of CBST scientists, guest experts in biophotonics, undergraduate interns, and educators. Participants selected what they thought was the best answer, provided a relevance and a difficulty rating (on a 1-3 scale), and gave written comments about item quality for each of the twenty questions.

The data collected was analyzed and used to calculate alpha reliability[1], assess concurrent validity, eliminate poorly worded or ambiguous items, and identify the most relevant concepts. In summary, the analysis provided evidence that although the knowledge instrument had acceptable internal consistency (α=.70) and was able to distinguish between biophotonics experts and non experts, there were a number of poorly worded items and questions that focused on memorized facts rather than conceptual understanding (e.g., specific absorption wavelength ranges). After reviewing the results, the internal evaluator recommended revising the instrument into a Biophotonics Concept Inventory (BPC). A concept inventory focuses on major ideas critical to understanding a particular field or phenomenon (e.g., force and motion) with items/ and responses that address common misconceptions found in the literature[2]. A number of concept inventories for use in undergraduate course assessment have been developed to date and include topics relevant to biophotonics such as the Light and Spectroscopy Concept Inventory designed for introductory astronomy courses[3]. It was also suggested that items would be more comprehensible to diverse audiences if visual representations were included whenever possible.

BPC development included several iterative steps: identification of major biophotonics concepts in consultation with CBST scientists and educators, a literature review of misconceptions about light and light/matter interactions, item formulation, and design of visual representations. The initial pilot version consisted of thirteen items that were administered at the 2008 CBST retreat. Participant comments were overwhelmingly supportive of the new approach and format. Item analysis led to selection of nine items for the program pilot BPC. An additional item was written afterwards and four items from the original biophotonics knowledge measure were added, resulting in a 14-item instrument for piloting in CBST programs.

The data and results reported in this paper are based on the program pilot BPC. The instrument was administered from 2009-2010 as a pre/post measure in all CBST programs with at least 8 hours of instruction in biophotonics. These include programs for undergraduate and high school students and middle/high school teachers. Each item response was coded as correct or incorrect. Scale validity (factor analysis), internal consistency (Cronbach's alpha), and mean comparisons (t-tests) were performed using SPSS. Recommendations for further BPC development are based on interpretation of program piloting results.

3. DATA

In addition to detailing the data for instrument analysis and program evaluation, this section also includes a description of sources used for identifying major concepts and misconceptions about light and light/matter interactions that underlie BPC item construction.

3.1 Basic concepts and common misconceptions about light and light/matter interactions

A small set of basic concepts underlying biophotonics were identified from three primary data sources: feedback from CBST scientists on items piloted at the CBST annual retreats; consultation with several individual CBST scientists; and review of science content standards. The standards documents included in our review were the National Science Education Standards[4], Benchmarks for Science Literacy[5], and the California State Science Content Standards[6]. The science literacy maps[7] associated with the Benchmarks for Science Literacy were especially helpful for identifying a small set of interconnected ideas critical for a basic understanding of light and light/matter interactions.

Researchers have found a number of common misconceptions about light that can be built into item responses as distracters in a conceptual inventory. Consequently, the inventory can be used as a diagnostic tool by examining incorrect response patterns to identify misconceptions that persist and plan instruction accordingly [8]. Guesne's work was especially helpful for identifying misconceptions about light [9] and Anderson and Smith's research provided insight into elementary students' naïve conceptions about the light-matter interactions involved in vision and color perception [10].

3.2 Scale components and internal consistency

The data for factor and reliability analysis were obtained from pre and post BPC questionnaires administered during 2009-2010 to participants in CBST middle school, high school, undergraduate, and teacher education programs that included more than eight hours of instruction in biophotonics. The number of BPC questionnaires completed in each group is shown in Table 1.

Table 1. Number of questionnaires from each education program included in factor and reliability analysis

Program group	Number of BPC questionnaires (n)
Middle School	44
High School	51
Undergraduate	42
Grades 5-12 teachers	12

3.3 Means comparisons and program evaluation

The data described in section 3.2 (see Table 1) also was used for mean comparisons between groups as an indicator of concurrent validity. In addition, effectiveness as an evaluation measure was checked by administering the BPC to participants in a 2010 undergraduate biophotonics course at University of California, Davis (n=20) at the beginning and end of the course and analyzed for improvement in scores and effect size.

4. Results

3.1 Major concepts and common misconceptions about light and light/matter interactions

The items in the BPC address the overall idea that light energy has predictable properties when it interacts with matter and, specifically, the following four major concepts included in the Benchmarks for Science Literacy (SMS-BMK) [11] and the National Science Standards (NSES) [12]:

1a. Electromagnetic waves result when a charged object is accelerated or decelerated. Electromagnetic waves include radio waves (the longest wavelength), microwaves, infrared radiation (radiant heat), visible light, ultraviolet radiation, x-rays, and gamma rays. The energy of electromagnetic waves is carried in packets whose magnitude is inversely proportional to the wavelength. (NSES)

1b. A great variety of radiations are electromagnetic waves...Their wavelengths vary from radio waves, the longest, to gamma rays, the shortest. In empty space, all electromagnetic waves move at the same speed-the "speed of light." *4F/H3bcd* (ID: SMS-BMK-1779)

2. Light from the sun is made up of a mixture of many different colors of light, even though to the eye the light looks almost white. Other things that give off or reflect light have a different mix of colors. *4F/M1* (ID: SMS-BMK-0217)

3a. Something can be "seen" when light waves emitted or reflected by it enter the eye. *4F/M2* (ID: SMS-BMK-0218)

3b. To see an object, light from that object--emitted by or scattered from it--must enter the eye. (NSES)

4a. Light travels and tends to maintain its direction of motion until it interacts with an object or material. Light can be absorbed, redirected, bounced back, or allowed to pass through. *NEW BENCHMARK* (ID: SMS-BMK-1829)

4b. Light interacts with matter by transmission (including refraction), absorption, or scattering (including reflection). (NSES)

Items and responses were selected/written to address common misconceptions identified for light, vision, and color [13]. While there has been research into misconceptions that children and adults have about light and optics [14], much has taken place in the context of physics [15] or astronomy [16]. Little work has been done specifically looking into misconceptions related to light-tissue interactions, the foundation of biophotonics. Consequently, BPC items that addressed interactions between light and living materials were based on misconceptions observed through developers' experiences working with students rather than empirical evidence. Given that tissue optics "occupies the middle ground between physics and biology" [17], science education would benefit from misconception studies in biophtonics.

BPC items cannot be shared in this paper because they are used for program evaluation and cannot be released. The general concept associated with each item is substituted for the actual item itself.

4.2 Scale components and internal consistency

Principal component analysis was performed in SPSS to see how the items worked together to measure biophotonics concepts. To determine the appropriateness of factor analysis, the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy and Barlett's test of sphericity were performed. The underlying assumption is that items are correlated because they are measuring the same thing. If the items are not correlated then there is no factorability—each item is unique onto itself. The KMO= .637 which means that the factors extracted will account for a fair, but not substantial, amount of the variance (Hair *et al.*, 1998). Bartlett's test of sphericity tests the null hypothesis that the variables in the population correlation matrix are uncorrelated. The significance level was .0000—small enough to reject the hypothesis. The KMO and Bartlett's test results indicate the sample size is adequate and items are sufficiently correlated to warrant factor analysis.

Table 2. Rotated Component Matrix[a]

	Component			
	1	2	3	4
1. Color seen when mixing light				.602 (2)
2. Need light to see	.413 (3)			
3. Color seen with yellow light from sodium lamp on blue object	.323		.422 (2)	
4. What orange filter does to white light			.488 (2)	
6. Color light for plant to grow the best.		.703 (4)		
7. Wavelength lower energy than blue	.536 (1)	.419		
8. Green-fluorescing protein molecule exposed to blue light	.554 (4)			
9. What infrared camera measures	.798 (1)			
10. Interaction between green laser and human finger			.656 (4)	
11. What cancer is			.707	
12. Color vision absorption peaks		.715 (4)		
13. How light travels in a constant medium				.784 (4)
14. Characteristics of ultraviolet light	.454			.465 (1)
Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.				
a. Rotation converged in 6 iterations.				

Principal components analysis with varimax rotation resulted in five factors and explained a moderate amount (55%) of the variance. The five-component extraction indicates how items clustered together. Recommendations for best fitting to data with factor analysis include the following: item loadings above .32, no or few item cross loadings (an item that loads above .32 on two or more factors), and no factors with fewer than three items [18]. After applying these criteria to the initial rotated component matrix, one item (#5) was eliminated. The resulting rotated component matrix extracted four factors, explaining 49% of the variance (see Table 2). The factor structure is still not very clean, however. Although all items load onto a factor above .32, there are still three items that cross load (items #3, #7, and #14), and component #2 only has two items loading at their highest value. These results indicate that there are some remaining problems with test construction that may be improved upon by increasing sample size, adding additional items, and/or revising or removing items.

The interpretation of the components extracted is largely determined on a theoretical or conceptual basis [20]. In Table 2, the number in parenthesis after each item's highest factor loading represents which one of the four major concepts from the selected standards (described earlier) the item primarily addresses. One way of explaining the components would be if factors corresponded to one of the standards. However, although there are some promising patterns, the items in a component do not represent a single standard. For example, both items in component two address standard 4 (light-matter interactions) but there are also items that address standard number 4 in components three and four. Again, further review of items, adding more items for each standard, and a larger sample size may provide more favorable results. However, there is evidence that BPC items can be grouped into different clusters that likely have a conceptual basis.

Another tool to assess the quality of instrument construction instrument is Cronbach's alpha—a measure of internal consistency or reliability. Specifically, Cronbach's alpha measures how closely related a set of items are as a group. Generally, a Cronbach's alpha of .70 or above is desirable. However, the value of Cronbach's alpha is influenced by the number of items. Consequently, scales with a small number of items may still be reliable yet have alpha levels below the 0.7 standard. Cronbach did provide a formula to correct for alpha's sensitivity to the item number—an estimate of the mean inter-item correlation (ρ)—that is independent of scale length:

$$\rho = \frac{\alpha}{n - (n-1)\,\alpha}$$

where ρ = an estimator of reliability independent of scale length, α = coefficient alpha, and n = the number of items in the scale. Mean inter-item correlations between .15 and .20 are considered adequate for outcome measures [12]. The four factors extracted from the principal components analysis each have a mean inter-item correlation within that range, indicating that the factor scales can be considered reliable.

Table 3. Reliability estimates

Factor	Number of items	Cronbach's Alpha	Mean inter-item correlation (ρ)
All items	13	.628	.115
Factor 1	4	.498	.199
Factor 2	2	.424	.269
Factor 3	4	.423	.244
Factor 4	3	.417	.192

In summary, although principal component analysis yielded puzzling results for conceptually explaining item clustering, given the small number of items, the components and reliability analysis support retaining the remaining thirteen items, developing additional items to ensure that there are at least four items for each of the four standards, and further pilot testing. The remaining thirteen BPC items were used to examine means comparison for concurrent validity and program evaluation.

4.3 Means comparisons and program evaluation

One of the most important criteria for judging the quality of a measure, like the BPC, is the extent to which it measures what it claims to measure—it's validity. Content validity represents the possible range of items that the instrument should include. As described in the methodology section, CBST scientists conducted an expert review of the initial item pool to establish content validity for items included on the pilot BPC. Reviews by CBST retreat participant also helped establish face validity—a consensus that the items measure the intended construct(s).

Concurrent validity is another type of validity that establishes the extent to which the items can distinguish between experts and non-experts. The pilot BPC data administered to program participants included subjects with a range of expertise: gifted middle school students, students from a low-performing high school, undergraduates enrolled in a biophotonics course, and grades 7-12 teachers. A means comparison analysis (ANOVA) was conducted to investigate group differences on the BPC for all thirteen items combined (total percent) and the four factors extracted from the components analysis. The results are reported in Table 4.

Table 4. ANOVA BPC pilot group comparisons

		N	Mean	Std. Deviation	ANOVA	Sum of squares	df	F	Sig
Total percent correct	MS	44	.5903	.16117	Between Groups	1.313	3	15.456	.000
	HS	50	.5192	.17960	Within Groups	4.077	144		
	UG	42	.7414	.16992	Total	5.389	147		
	teacher	12	.7308	.13323					
	Total	148	.6206	.19147					
Total percent correct Factor 1	MS	44	.6477	.24894	Between Groups	2.261	3	11.575	.000
	HS	50	.4900	.28997	Within Groups	9.376	144		
	UG	42	.7679	.23680	Total	11.636	147		
	teacher	12	.8333	.16283					
	Total	148	.6436	.28135					
Total percent correct Factor 2	MS	44	.1136	.23781	Between Groups	7.382	3	30.202	.000
	HS	50	.1400	.22678	Within Groups	11.732	144		
	UG	42	.6310	.36725	Total	19.113	147		
	teacher	12	.2500	.33710					
	Total	148	.2804	.36059					
Total percent correct Factor 3	MS	44	.7330	.25517	Between Groups	.677	3	3.277	.023
	HS	50	.6200	.30822	Within Groups	9.912	144		
	UG	42	.7798	.22226	Total	10.589	147		
	teacher	12	.7708	.19824					
	Total	148	.7111	.26839					
Total percent correct Factor 4	MS	44	.6439	.36229	Between Groups	.491	3	1.698	.170
	HS	50	.6733	.28164	Within Groups	13.882	144		
	UG	42	.7183	.30463	Total	14.373	147		
	teacher	12	.8611	.22285					
	Total	148	.6926	.31269					

It is important to note that mean differences were not expected between the middle and high school students but between the middle/high school students, undergraduates students/ teachers with higher scores expected for the latter. There is a significant difference between groups at $p<.05$ level for the total and all factors except factor four. However, values for factor 4 trend in the expected direction. The means for Factor 2 are quite low for middle/high school students and teachers in comparison with the undergraduates. A closer examination of items grouped under Factor 2 indicate that item #10 (interaction of green laser with human finger) has lower mean scores across groups (0.42 - 0.66) than other items and should be further investigated by interviewing a sample of participants to see how they interpreted the drawings included in that item.

In general, the ANOVA results support the concurrent validity of the BPC. For most factors, participant groups with less expertise tend to, on average, score lower on items than those with more advanced education.

The primary purpose of the BPC is to serve as a pre and post measure for evaluating program effects on participant understanding of biophotonics. To investigate it's effectiveness as an evaluation tool, the BPC was administered before and after instruction to undergraduates from a variety of majors who enrolled in a 2010 biophotonics course offered as part of an interdisciplinary studies program at University of California, Davis. The expectation was that scores would improve as a result of instruction. The results of the student's t-test comparison for all items combined and the four factors are reported in Table 5.

Table 5. Paired sample test pre and post BPC scores

		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Total post - Total pre	.16346	.13669	.02790	.10574	.22118	5.858	23	.000
Pair 2	Total post Factor 1 - Total pre Factor 1	.26042	.20161	.04115	.17528	.34555	6.328	23	.000
Pair 3	Total post Factor 2 - Total pre Factor 2	.16667	.50361	.10280	-.04599	.37932	1.621	23	.119
Pair 4	Total post Factor 3 - Total pre Factor 3	.12500	.19505	.03981	.04264	.20736	3.140	23	.005
Pair 5	Total post Factor 4 - Total pre Factor 4	.08333	.17720	.03617	.00851	.15816	2.304	23	.031

BPC scores were, on average, significantly higher ($p<.05$) after instruction for all items combined (total percent) and for each factor except factor two although there was a positive mean difference for factor two as well. The results indicate that the items on the BPC are sensitive enough to detect changes in understanding of concepts related to biophotonics as a result of program instruction.

However, significance alone does not mean that the difference matters in a practical sense. As Thalheimer and Cook point out, "Whereas statistical tests of significance tell us the likelihood that experimental results differ from chance expectations, effect-size measurements tell us the relative magnitude of the experimental treatment" [22]. Effect size is the difference between the control and treatment means divided by the standard deviation. Because effect sizes take the standard deviation into account, the magnitude of the effect can be compared across experiments. Cohen's *d* is the recommended effect size statistic [23]. However, in the case of repeated measures, as with pre/post BPC testing, pre and post scores are likely correlated. The repeated measures Cohen's d is based on the pooled standard deviation corrected for the correlation [24]. Table 6 reports the effect sizes for BPC items combined and disaggregated into the four factors identified earlier.

Table 6. Effect size (repeated measures Cohen's d)

	Mean	SD	N	Mean Dif	Pooled SD	Correlation	Cohen's d	Interpretation
Post –all items	0.91	0.09	24	0.16	0.14	0.49	1.68	Large positive difference
Pre-all items	0.75	0.16	24					
Post - F1	0.99	0.05	24					Large positive
Pre - F1	0.73	0.22	24	0.26	0.20	0.46	1.78	difference
Post-F2	0.75	0.36	24					Large positive
Pre-F2	0.58	0.35	24	0.17	0.50	0.99	1.05	difference
Post-F3	0.95	0.13	24					Large positive
Pre-F3	0.82	0.19	24	0.13	0.19	0.28	0.8	difference
Post-F4	0.88	0.22	24					Large positive
Pre-F4	0.79	0.24	24	0.08	0.18	0.70	0.8	difference

Cohen's *d* values above 0.80 are considered evidence of a large difference [25]. Effect size measures for all items combined and each of the factors are greater than .08 and indicate that the IST8A class is an intervention that has a large positive impact on BPC scores. In other words, the mean difference between BPC scores before and after the course is not only significant but the magnitude of the effect is large.

5. Conclusion

Although the BPC represents a modest effort towards creating a conceptual inventory for use in biophotonics course and program evaluation, it is a promising tool. The inclusion of CBST scientists and program participants provided the necessary expertise and sample sizes to investigate measurement qualities including instrument scales, validity, and reliability. Involving a science center (or department, division, etc.) community in developing a conceptual inventory affords opportunities for rich discussion and consensus building around major concepts that undergird a discipline. This is especially important as cutting-edge interdisciplinary optics fields, like biophotonics, become more the rule, rather than the exception, for scientific research.

While the BPC still needs improvement, analysis of pilot data demonstrates adequate validity and reliability to use an evaluation measure. The measure would be improved by adding additional items and an *a priori* structure that is conceptually stronger (e.g., sufficient items linked to a small set of content standards) and could be confirmed through factor analysis. The biophotonics education community would benefit from a conceptual inventory similar to those available for astronomy and physics. Developing a higher quality instrument presents an interesting and useful joint project for SPIE, and other, groups interested in photonics education.

REFERENCES

[1] Cronbach, L. J., *Essentials of psychological testing* (3rd ed.), New York: Harper & Row (1970).

[2] Evans, D.L., Gray, G.L., Krause, S., Martin, J., Midkiff, C., Notaros, B.M., Pavelich, M., Rancour, D., Reed-Rhoads, T., Steif, P., Streveler, R., Wage, K., "Progress on concept inventory assessment tools," 33rd Annual Frontiers in Education (FIE'03), Volume 1, T4G-8 (2003).

[3] Bardar, E.M., Prather, E.E., Brecher, K. and Slater, T. F., "Development and evaluation of the light and spectroscopy concept inventory," Astronomy Education Review 5(2), 103-113 (2006).

[4] http://www.nsta.org/publications/nses.aspx

[5] http://www.project2061.org/publications/bsl/default.htm

[6] www.cde.ca.gov/be/st/ss/documents/**science**stnd.pdf

[7] http://strandmaps.nsdl.org/

[8] Bardar et al. (2006).

[9] Guesne, E., "Light." In Driver, R. (Ed.), [Children's ideas in science], Open University Press, Pennsylvania, 10-32 (1985).

[10] Anderson, C.W., and Smith, E.L., "Children's conception of light and color: Understanding the role of unseen rays," Institute for Research on Teaching, Michigan State University, (1986).

[11] Benchmarks for science literacy http://www.project2061.org/publications/bsl/default.htm

[12] National science education standards http://www.nap.edu/openbook.php?record_id=4962

[13] Bardar et al. (2006), Anderson and Smith (1985), Guesne, E. (1985)

[14] Pompea1, S.M. Dokter, E.F., Walker, C.E., and Sparks, R.T., "Using misconceptions research in the design of optics instructional materials and teacher professional development programs," SPIE (2007).

[15] Reiner, M., Slotta, J.D., Chi, M., Resnick, L., "Naive physics reasoning: A commitment to substance-based conceptions," Cognition and Instruction 18(1), 1-34 (2000).

[16] Bardar et al (2006).

[17] Jacques, S.L., "Special tutorial: Tissue optics and laser-tissue interactions, 2010 CBST Retreat.

[18] Costello, J.B., & Osborne, J.W., "Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis," Practical Assessment, Research and Evaluation 10(7) (2005) http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.110.9154&rep=rep1&type=pdf.

[19] Costello & Osborne (2005).

[20] Morgan, G.A., & Griego, O.V., Factor analysis: Data reduction with principal components analysis, [Easy use and interpretation of SPSS for windows: Answering research questions with statistics], 111-124 (1998).

[21] Spiliotopoulou, G., "Reliability reconsidered: Cronbach's alpha and paediatric assessment in occupational therapy," Australian Occupational Therapy Journal 56, 150–155 (2009).

[22] http://www.work-learning.com/effect_sizes.htm

[23] Cohen, J., [Statistical power analysis for the behavioral sciences (2nd ed.)]. Hillsdale, NJ: Lawrence Earlbaum Associates (1988).

[24] Rosnow, R. L., & Rosenthal, R., "Computing contrasts, effect sizes, and counternulls on other people's published data: General procedures for research consumers," Psychological Methods, 1, 331-340 (1996).

[25] Cohen, J. (1988).