Review Article

The affect of symbols: creating the graphic differential through synesthesia and metaphor

Abstract

Previous research with the semantic differential technique has established the cross-cultural generality of three basic affective dimensions identified as evaluation, potency, and activity. The purpose of this study is to develop a graphic equivalent of the pan-cultural short form semantic differential. This graphic differential is composed of bipolar scales made up of pictographic opposites on which subjects rate a concept by placing a check mark on the seven-point scale. Such a non-verbal instrument would be culture free and would have certain distinct advantages of the verbal semantic differential form. Data are reported for a pre-test of the graphic instrument in five language-culture communities (American English, Delhi Hindi, Finnish, German, and Japanese). The study succeeds in identifying several graphic scales that provide reliable and effective indices of the evaluation dimension of affect. These scales exhibit and conform to the symmetry law of Gestalt and could be used as a culture free attitude measure. With respect to the potency and activity dimensions, the results are less satisfactory, although some of the graphic scales can be used for this purpose when chosen indigenously (i.e. independently for each culture). Evidence of visual denotative contamination is found and discussed.

Keywords: symbols, affect, semantic differential, graphic pictographs, cross-cultural, synesthesia, metaphor, art

Leon James
Department of Psychology, University of Hawaii, USA

Correspondence: Leon James, Department of Psychology, University of Hawaii, 2430 Campus Road Honolulu, HI 96822, USA, Email leon@hawaii.edu

Introduction

The rationale that motivated this study presupposes two kinds of universal tendencies. The first concerns the existence of three pan-human affective dimensions identified as Evaluation, Potency, and Activity. The second relates to the assumption that elementary visual forms of an abstract nature evoke affective-like reactions that are not dependent on culturally variable and arbitrary meanings; instead, these reactions are presumed to be mediated by pan-human synesthetic tendencies in response to the inherent visual properties of the stimulus forms. I shall now present in brief form the arguments that convinced me of the acceptability of these two presuppositions.

The Pan-cultural semantic differential

The early work of Osgood and his associates on the semantic differential technique, as summarized in Osgood et al.[1] clearly established the following fact about the modes of verbal qualification in the usage of English by Americans: if people are asked to rate concepts on bi-polar scales made up of a large number of randomly selected adjectival opposites (e.g. good-bad, sweet-sour, serious-fickle, deep-shallow, young-old, etc.), a substantial portion of the variability in ratings on the various scales (between 60% and 65%) can be accounted for by only three factors. This remarkable finding represents at once the strength and the weakness of the semantic differential technique. It is a weakness, indeed a fatality, for its use as a general measure of the meaning of words since it is obvious that in order to differentiate the many thousands of words that the ordinary speaker of English knows, many more dimensions than three will be needed. It is a strength, for its use as a measure of the affect of words since the small number of factors (only three) insures their highly generalized nature so that the rating instrument gains increased usefulness--and indeed, it has been successfully used with all categories of words and with non-verbal stimuli (geometric forms, color patches, works of art, sounds, objects, people, etc.; see the bibliography in Snider and Osgood.[2]

Having established the American English generality Or the three-factor affective system and having identified the three factors as Evaluation (E), Potency (P), and Activity (A), the next step was to inquire whether this specific language culture generality extended to other language/culture communities. The affirmative answer which was fully expected ran contrary to a broad interpretation Or the Sapir-Whorf Relativity Hypothesis-- and the empirical work which ensued left no doubt as to which hypothesis was correct: each language/culture community, up to 28 in all, exhibited the same pattern.[3,4] The E-P-A factor system turned up in every case by independently replicated indigenous (i.e. not by translation) elicitation procedures followed in a carefully standardized method in each location.

One of the useful tools that emerged from this cross-cultural work in the "pan-cultural short form semantic differential" which consists of 12 bi-polar adjectival scales--four scales for each of the three universal affective factors--with each of the 28 communities possessing a version in their own language. The advantage of this tool is that the E-P-A value of stimuli can now be directly compared across the various cultures. For example, in the World Atlas of Semantic Profiles.[5] comparisons are made on the affective meaning of 600 concepts as rated by unilingual speakers in each location. These comparisons lead to hypotheses about cultural universals as well as differences and peculiarities proper to a cultural group (for examples, see 3).

The cross-cultural generality of visual-verbal synesthetic tendencies

A demonstration I like to try when discussing with my students the general problem of metaphor goes as follows. I ask them to imagine a collection of stimuli and group them into sets that "fit together"; the collection is made up of various sensory stimuli in different modalities: the loud clang of a bell, the feel of sandpaper, an itchy scratch, a white circle, a straight line, the taste of chocolate, a rising

melody, various colors, and so in. Invariably, there is good agreement on how the sub-sets fit together: there is a set that contains a white circle, a straight line, a caressing touch, a sweet taste, a bright color, a rising melody: there is another set containing a loud clang, saturated colors, deep low sounds, hard surfaces, and irritating scratches; and a third set is made up of bright colors, wavy lines, rhythmic sounds, and flashes of light.

What is the basis of these agreements? Osgood's explanation, in summary form, runs as follows: "...the highly generalized nature of the affective reaction system--the fact that it is independent of any particular sensory modality, yet participates with all of them--is... the psychological basis for synesthesia and metaphor. It is because such diverse sensory experiences as a white circle (rather than black), a straight line (rather than crooked), a rising melody (rather than a falling one), a sweet taste (rather than a sour one), a caressing touch (rather than an irritating scratch) can all share a common affective meaning that one can easily and lawfully translate from one modality into another in synesthesia and metaphor...I would suggest that this meaning system is intimately related to the nonspecific projection systems from the hypothalamic, reticular, or limbic systems and their cortical connections in the frontal lobes--both are gross, nondiscriminative, but highly generalizable systems and both are associated with the affective, purposive and motivational dynamics of the organism".[6 p.21]

Is it just chance, or universal affect, that is at work in our linguistic metaphors when we say that we move up in society when we better our social standing, when we form an inner circle of friends to keep strangers out, when we refer to happy faces as bright, when our lot is sweet when we are lucky, when a pleasant voice is caressing, and so on? There is evidence.[7] that the up-down and in-out analogies of social standing appear in other languages as well, and the "laws of good Gestalt" which Wertheimer and Kohler talked about were certainly meant to be a universal property of the human visual system. Osgood's.[8] study on the cross-cultural generality of certain visual verbal synesthetic tendencies demonstrated good agreement on the part of subjects drawn from three different linguistic/cultural backgrounds: American college students, Japanese foreign students in the US, Mexican-Spanish speaking persons in the Southwest, and Navajo Indians.

These subjects were shown cards on which were drawn two simple figures: e.g. a white and a black circle, two arrows (one pointing upward, the other downward), two lines (one thick, the other thin), two circles (one large, the other small), etc. subjects were given words in their language, one at a time (e.g. heavy, Man, yellow, sad, EXCITEMENT) and they were to point to one or the other of the visual alternatives on the card according to what "seemed most appropriate to that concept." Significant and substantial agreements were found for most of the answers both intra-culturally and cross-culturally.

The cross-cultural graphic differential

The preceding discussion tried to establish the two presuppositions mentioned in the Introduction as they constitute the rationale for attempting to construct a "graphic differential" (henceforth GD). If one is convinced that the E-P-A affective system is universal and that it mediates the synesthetic reactions to visual stimuli, then it makes sense to attempt to construct a visual equivalent of the pan-cultural short form semantic differential (the latter will henceforth be referred to as SD). Thus, the GD was to become a visual equivalent of the SD in all its respects: it should be made up of bi-polar terms (henceforth to be referred to as "pictographic opposites"); these should have the

property of a one-dimensional scale with seven polarity positions including a mid-point; there should be a small number of them and of a size suitable to fit a single page.

There are a number of reasons why such an instrument would be desirable. The following appear to me to be the most salient:

It would represent an alternative form to the present SD: Several advantages are to be had by this. Some of our colleagues are bothered by the fact that the various indigenous forms of the SD are not translation equivalents of each other, yet we infer that they all measure the same affect. They might find the GD more comfortable to work with since the forms for all languages will now be (visually) identical. Other colleagues have worried about the effects on SD ratings in situations where the experimental design called for repeated testing; the availability of an alternative non-verbal form should relieve some of this anxiety on their part. In some situations, verbal written forms of the SD are inappropriate; e.g. with pre-school children, with illiterate adults or those from a pre-literate society, with aphasics who have lost the ability to read the GD, should it prove practicable with these subjects on other grounds (e.g. the concept of a scale, the rating task), would then allow for applications now impossible with the SD.

It would provide insights into the nature of synesthesia and metaphor: Arguments to this effect have already been presented above. Osgood's neurophysiological speculations about synesthesia, while just that, nevertheless provide a rationale for interpreting a large variety of disparate phenomena.

The specific visual characteristics of the pictographic scales that have cross-culturally similar affect-- should such an outcome turn out to be feasible--would provide specific hypotheses about the nature of the human visual system and its dynamic relation to the human motivational system. Such an outcome would also be relevant to certain semantic universals that arc grounded in the perceptual system such as visual metaphors in language.

It would contribute to the psychology of art and aesthetics: A demonstration of the existence of visual forms that have an unlearned, universal meaning (of the affective sort) would contribute to the long standing controversy between critics who maintain an "absolutist" position vis a vis "good" art and proponents of the "relativist" school. There is keen interest today among graphic designers in constructing international symbols to denote public facilities and services; a cross-cultural GD should prove useful in identifying characteristics of visual symbols that carry "emotional tones" acceptable in all cultures.

Method

The design of this study was straightforward. The plan consisted of requiring subjects to rate 50 concepts on a large number of GD scales as well as on the 12 verbal scales of the SD. These ratings would then be intercorrelated and factor analyzed so QS to identify those GD scales that acted most like the E-P-A scales of the SD instrument. The factor structure, the size of the loadings of their defining scales, and intercorrelations between GD and SD composite scores would form the criteria by which we could evaluate the existence of cross-culturally valid pictographic scales.

The test booklets

The selection of GD scales was done on a purely intuitive basis. We asked a number of our cross-cultural colleagues to send us simple abstract designs which we either replicated in their original form or adapted to the designs we made up. We retained a total of 64 pairs and these are reproduced here as Figure 1. The size of the squares

enclosing each pictograph was 25mm x 25mm, so that eight scales could fit on a standard 8" x 11" page. The designs were originally drawn with black ink in large size on individual cards then reduced to the final size through photographic process. The left-right orientation of each pair of pictographic alternatives was determined by a random process. Stencils were cut through an electronic process and run off on a standard Gestetner mimeograph machine. The mimeographed sheets were then mailed to each field location where they were assembled in booklet form after the 50 concepts were added (written in, typed, or printed) in the appropriate language. These concepts were randomly chosen from a set of 100 words previously used in the work on the development of the pan cultural SD and had already been translated in the various languages. The words in English are listed in Table 1.
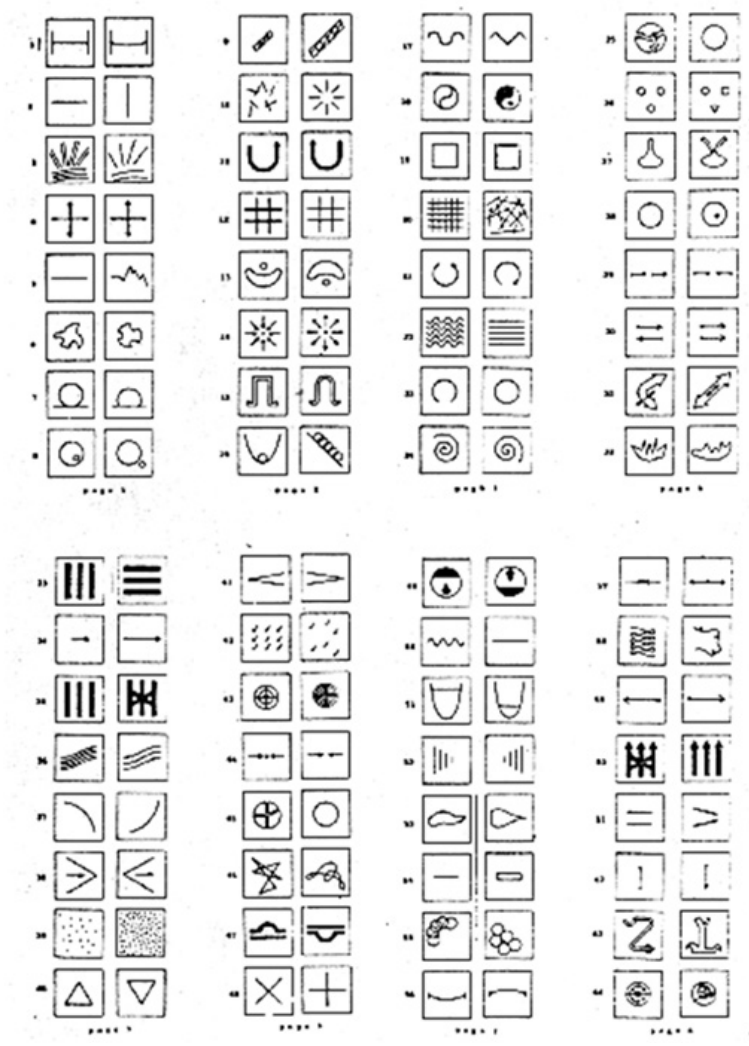
Figure 1 The Pictographic Opposites in the Graphic Differential.

Table 1 The fifty stimulus concepts used

Group 1	Group 2	Group 3	Group 4	Group 5
1. Love	11. Husband	21. Tooth	31. Wind	41. Color
2. Progress	12. Guilt	22. Freedom	32. Hand	42. Belief
3. Woman	13. Knot	23. Sun	33. Laughter	43. Head
4. Star	14. Story	24. Success	34. Smoke	44. Battle
5. Food	15. Mother	25. Thief	35. Father	45. Respect
6. Peace	16. Death	26. Work	36. Hope	46. Life
7. Girl	17. Chair	27. Man	37. Root	47. Snake
8. Anger	18. Horse	28. Defeat	38. Fear	48. Money
9. Tongue	19. Sleep	29. Cloud	39. Truth	49. Author
10. Crime	20. Trust	30. Policeman	40. Purpose	50. Punishment

Each field location received a standard set of instructions which were to be followed in the execution of the task. There were to be three booklets and two testing sessions for each subject. The first booklet was to consist of 48 pages as follows: six concepts were to be rated on all 64 GD scales, there being eight scales per page; thus, each concept appeared on top of eight successive pages, The second booklet U85 made up of 32 pages: four concepts each being rated on the 64 GD scales. The third booklet consisted of the appropriate SD scales in each language. There were 12 scales per page with the concept to be rated appearing on top. These consisted of the ten concepts of booklets 1 and 2. The six concepts for booklet 1 were to be chosen at random and half the booklets for each group were to be assembled in one random order and the other half in the reverse order. Similar random procedures were used for booklets 2 and 3.

Each subject rated all three booklets: booklet 1 during session one and booklets 2 and 3 (in that order) during session two, The testing was done in groups, there being at least 2325 subjects in each group. Since each subject rated only 10 concepts, five groups were needed for the 50 concepts. Each session lasted about an hour. The subjects in each location were drawn from a school population of young men and women within the age range of 17 to 21 (in the US these were college freshmen and sophomores).

Instructions

Written instructions appeared in the front of each booklet. These were adapted from the usual instructions with the semantic differential technique.[1] Clarifications about the task were given orally by the experimenter in charge whenever these seemed desirable.

The participating sites

In the selection of sites for data collection an attempt was made to diversify as much as possible both culture and language, at least, as much as our network of 28 cooperating language/culture communities allowed for. The present report summarizes the findings for the following five locations: American-English (AE; Urbana, Illinois), Delhi-Hindi (DH; India), Finland-Finnish (FF; Helsinki), Germany-German (GG; Munster, Westfalia), and Japan-Japanese (JP; Tokyo). The pan cultural SD scales for each language are given in their English (approximate) translations in Table 2. The numbers 55-7 give the scale identification for subsequent discussion. The letters in parenthesis preceding each scale refers to the affective dimension it taps: E = Evaluation, P = Potency, and A = Activity. The scales are oriented in the random direction presented in the test booklets. For each scale, that qualifier is underlined which represents the positive aspect of the factor keyed towards the American-English scales (viz. for E: helpful', nice, good, sweet; for P: big, strong, powerful, deep; for A: young, fast, alive, noisy).

Table 2 Pan-Cultural short form semantic differential (sd) for five languages in their approximate English translations

Scale I.D.	American-English
65.	(E) helpful-unhelpful
66.	(P) big-small
67.	(A) old-young
68.	(A) slow-fast
69.	(E) nice-awful
70.	(P) strong-weak
71.	(E) good-bad
72.	(A) alive-dead
73.	(P) powerless-powerful
74.	(A) quiet-noisy
75.	(P) deep-shallow
76.	(E) sour-sweet

Scale I.D.	Delhi-Hindi
65.	(P) cowardly-brave
66.	(A) hard-soft
67.	(P) fast-slow
68.	(E) nectarlike-poisonous
69.	(A) serious-fickle
70.	(A) thin-thick
71.	(P) heavy-light
72.	(E) good-bad
73.	(P) easy-difficult
74.	(E) angry-glad
75.	(E) superior-inferior
76.	(A) strong-imperfect

Scale I.D.	Finland-Finnish
65.	(A) rigid-flexible
66.	(E) good-bad
67.	(A) clumsy-agile
68.	(E) nice-not nice
69.	(A) subdued-lively
70.	(P) large-small
71.	(E light-gloomy
72.	(P) delicate-sturdy
73.	(A) fast-slow
74.	(P) weak-strong
75.	(E) pleasant-unpleasant
76.	(P) heavy-light

Scale I.D.	American-English
65.	(E) helpful-unhelpful
66.	(P) big-small
67.	(A) old-young
68.	(A) slow-fast
69.	(E) nice-awful
70.	(P) strong-weak
71.	(E) good-bad
72.	(A) alive-dead
73.	(P) powerless-powerful
74.	(A) quiet-noisy
75.	(P) deep-shallow
76.	(E) sour-sweet

Scale I.D.	Delhi-Hindi
65.	(P) cowardly-brave
66.	(A) hard-soft
67.	(P) fast-slow
68.	(E) nectarlike-poisonous
69.	(A) serious-fickle
70.	(A) thin-thick
71.	(P) heavy-light
72.	(E) good-bad
73.	(P) easy-difficult
74.	(E) angry-glad
75.	(E) superior-inferior
76.	(A) strong-imperfect

Scale I.D.	Finland-Finnish
65.	(A) rigid-flexible
66.	(E) good-bad
67.	(A) clumsy-agile
68.	(E) nice-not nice
69.	(A) subdued-lively
70.	(P) large-small
71.	(E) light-gloomy
72.	(P) delicate-sturdy
73.	(A) fast-slow
74.	(P) weak-strong
75.	(E) pleasant-unpleasant
76.	(P) heavy-light

Table 2 Continued...

Scale I.D.	American-English
65.	(E) helpful-unhelpful
66.	(P) big-small
67.	(A) old-young
68.	(A) slow-fast
69.	(E) nice-awful
70.	(P) strong-weak
71.	(E) good-bad
72.	(A) alive-dead
73.	(P) powerless-powerful
74.	(A) quiet-noisy
75.	(P) deep-shallow
76.	(E) sour-sweet
Scale I.D.	**Delhi-Hindi**
65.	(P) cowardly-brave
66.	(A) hard-soft
67.	(P) fast-slow
68.	(E) nectarlike-poisonous
69.	(A) serious-fickle
70.	(A) thin-thick
71.	(P) heavy-light
72.	(E) good-bad
73.	(P) easy-difficult
74.	(E) angry-glad
75.	(E) superior-inferior
76.	(A) strong-imperfect
Scale I.D.	**Finland-Finnish**
65.	(A) rigid-flexible
66.	(E) good-bad
67.	(A) clumsy-agile
68.	(E) nice-not nice
69.	(A) subdued-lively
70.	(P) large-small
71.	(E) light-gloomy
72.	(P) delicate-sturdy
73.	(A) fast-slow
74.	(P) weak-strong
75.	(E) pleasant-unpleasant
76.	(P) heavy-light
Scale I.D.	**Germany-German**
65.	(A) quiet-agitated
66.	(E) unpleasant-pleasant
67.	(P) small-big
68.	(A) lively-sedate
69.	(E) good-bad
70.	(P) heavy-light
71.	(A) still-noisy
72.	(E) unfriendly-friendly
73.	(P) tender-vigorous
74.	(A) fast-slow
75.	(E) beautiful-ugly
76.	(P) strong-weak
Scale I.D.	**Japan-Japanese**
65.	(A) active-inactive
66.	(P) small-big
67.	(P) heavy-light
68.	(E) good-bad
69.	(P) brave-cowardly
70.	(A) hard-soft
71.	(A) plain-colorful
72.	(E)comfortable-uncomfortable
73.	(E) happy-sad
74.	(A) noisy-quiet
75.	(E) pleasant-unpleasant
76.	(P) difficult-easy

Results

As indicated earlier, the basic analysis consists of factoring the intercorrelations between all the scales, selecting the highest loading GD scales that are equivalent to the SD E-P-A scales, and intercorrelating the resultant composite scores for the 50 concepts as the test for equivalence between the verbal and visual forms of the instrument. The results are presented in Tables 3 & 4.

Table 3 Scales with the highest loadings for the first three factors of the principle axes solution involving the 64 and the 12 SD scales.

American English

25.	-.94	.08	-.12	**15.**	-.40	.74	-.06	**9.**	-.14	-.24	.60
64.	.93	.12	.05	**46.**	-.19	.73	-.25	**12.**	.04	.49	-.57
19.	.92	.10	.07	**17.**	.50	-.67	.16	**39.**	.28	-.34	.57
31.	-.92	-.09	.03	**1.**	.45	.62	-.07	**3.**	-.18	.14	-.55
10.	-.91	-.09	.10	**6.**	-.53	-.62	-.09	**16.**	.21	-.18	.54
4.	.90	-.15	.14	**32.**	-.65	.52	-.21	**34.**	-.54	.01	.51
61.	.90	-.06	-.05	**44.**	-.02	.52	-.17	**37.**	-.38	.09	.45
28.	.89	-.19	.08	**22.**	-.65	-.54	-.12	**14.**	-.54	.43	.41
45.	-.89	.08	-.15	**50.**	-.64	-.48	-.30	**43.**	.45	.18	.40
71.	.89	-.13	.00	**66.**	.44	.28	-.49	**74.**	.36	-.27	.47
65.	-.88	.05	.03	**70.**	.69	.14	-.34	**68.**	.10	-.03	.33
69.	.88	-.20	-.01	**73.**	-.55	-.14	.50	**72.**	.64	-.15	-.21
76.	-.85	.25	.01	**75.**	.65	-.03	-.04	**67.**	-.28	.25	.06

Delhi-Hindi

22.	-.79	-.18	-.30	**1.**	.50	.47	.09	**9.**	.40	.05	.53
35.	.78	.06	.16	**2.**	.28	-.37	.27	**5.**	.57	.26	.49
10.	-.77	-.07	-.16	**34.**	-.30	-.38	.33	**50.**	-.62	-.33	-.45
48.	-.75	-.14	-.28	**50.**	-.62	-.33	-.45	**16.**	.31	.22	.44
25.	-.74	-.08	-.35	**32.**	-.21	.32	-.31	**53.**	-.65	-.26	-.41
4.	.72	.23	.31	**7.**	.58	.31	.28	**28.**	.52	.08	.40
62.	.72	-.01	-.07	**27.**	.54	.29	.30	**61.**	.67	.17	.39
64.	.72	.06	.25	**17.**	.09	-.28	.08	**39.**	.40	.15	.38
45.	-.71	-.11	-.17	**46.**	.33	.28	.03	**33.**	.20	-.14	-.38
72.	.88	-.07	.07	**71.**	.03	.62	-.24	**66.**	-.22	.57	-.27
68.	.84	-.16	.07	**67.**	-.10	.30	-.32	**70.**	.11	-.41	.18
74.	-.84	.15	-.08	**73.**	.05	-.16	.31	**69.**	.45	.49	-.07
75.	.82	-.08	.09	**65.**	-.66	-.04	.27	**76.**	.61	.27	.01

Finland/Finnish

19.	.93	.12	-.10	**15.**	-.10	.81	.08	**16.**	-.08	-.03	.64
25.	-.93	.01	.04	**17.**	.49	-.77	.06	**38.**	-.14	.42	.52
45.	-.91	.09	-.10	**44.**	-.11	.70	.02	**24.**	-.45	.09	.52
5.	.89	.04	.16	**46.**	-.21	.70	-.07	**9.**	.04	-.36	.51
10.	-.89	-.04	.17	**22.**	-.49	-.70	-.01	**2.**	.32	-.38	.49
7.	.88	.07	-.15	**7.**	-.17	.66	.24	**8.**	.35	.17	.48
4.	.87	.26	-.10	**32.**	-.59	.62	-.10	**59.**	-.46	.23	.46
20.	.87	.25	.04	**6.**	-.37	-.62	.05	**37.**	-.36	.16	.43
26.	.86	-.05	.03	**12.**	-.53	.58	.10	**34.**	-.64	-.06	.42
31.	-.86	-.17	.21	**1.**	.62	.56	-.09	**14.**	-.53	.55	.40
35.	.86	-.20	-.21	**50.**	-.62	-.54	-.07	**49.**	-.66	.01	.37
47.	.86	-.05	-.19	**58.**	.67	.50	.00	**60.**	-.78	.20	.37
66.	.90	-.10	-.11	**72.**	.20	-.77	.10	**73.**	-.12	-.29	-.44

Germany/German

64.	.91	.21	.07	**46**.	-.14	.83	-.21	**9**.	-.03	-.06	.66
10.	-.90	-.05	-.10	**1**.	.37	.82	-.18	**16**.	-.51	.38	.53
25.	-.90	.14	-.12	**22**.	-.41	-.81	.00	**33**.	.04	.40	-.62
31.	-.87	-.16	.12	**15**.	-.12	.85	-.19	**34**.	-.35	-.09	.57
4.	.84	.34	.07	**17**.	.27	-.79	.28	**2**.	-.11	-.35	.56
23.	-.84	-.18	-.06	**32**.	-.31	.75	-.37	**14**.	-.51	.38	.53
60.	-.84	.10	.26	**50**.	-.38	-.74	-.21	**56**.	-.28	-.27	.51
20.	.83	.36	.15	**6**.	-.41	-.72	-.11	**59**.	-.63	.15	.51
35.	.82	-.11	-.18	**12**.	-.07	.69	.07	**24**.	-.26	-.10	.45
45.	-.82	.11	-.11	**53**.	-.63	-.60	.12	**37**.	-.45	.10	.45
69.	.87	-.29	-.03	**70**.	-.19	.78	-.01	**65**.	.52	.11	.58
66.	-.82	.39	-.04	**73**.	.10	-.65	.39	**68**.	-.42	-.23	-.54
75.	.82	-.42	-.03	**76**.	.30	.42	-.37	**71**.	.28	-.21	.53

Japan/Japanese

25.	-.93	.11	-.09	**46**.	-.26	.84	-.12	**9**.	.06	-.20	.57
64.	.93	.14	-.04	**15**.	-.31	.81	-.09	**16**.	.48	.11	.55
45.	-.90	.11	-.06	**44**.	.00	.73	.12	**56**.	-.47	-.15	.55
26.	.89	-.12	.12	**17**.	.46	-.76	.13	**38**.	-.15	.44	.50
48.	-.89	.07	-.14	**1**.	.46	.70	-.12	**14**.	-.41	.41	.47
10.	-.88	-.17	.00	**6**.	-.44	-.70	.11	**2**.	.10	-.17	.44
23.	-.88	-.14	.01	**22**.	-.57	-.68	.12	**33**.	.10	.49	-.36
31.	-.88	-.06	.06	**32**.	-.56	.65	-.17	**34**.	-.51	-.09	.36
60.	-.88	.22	.17	**53**.	-.46	-.64	.18	**62**.	.79	-.04	-.36
40.	.87	.13	-.02	**50**.	-.60	-.60	.00	**12**.	.03	.56	-.35
68.	.91	-.09	-.06	**67**.	-.09	.76	-.04	**65**.	.49	-.05	-.62
75.	.89	-.23	-.03	**76**.	-.04	.60	-.09	**71**.	.07	.41	.43
72.	-.88	.21	.04	**66**.	-.32	-.19	.49	**70**.	-.08	.84	-.02

Table 4 Composite means for face validity concepts on the short form cross-cultural semantic (SD) and graphic (GD) differentials.

American-English

2. Progress	1.9	1.9	1.1	1.4	0.8	1.2
8. Anger	-1.6	0.6	0.5	-1.9	0.9	0.7
10. Crime	-2.3	0.8	-0.1	-1.6	1.0	0.4
22. Freedom	2.4	2.3	0.4	1.1	-0.3	1.1
24. Success	2.4	1.9	0.5	2.1	0.2	1.4
28. Defeat	-1.7	0.0	-.05	-1.9	-0.1	-1.1
29. Cloud	0.9	1.2	-0.4	0.0	-1.3	0.3
30. Policeman	0.9	1.2	0.4	0.5	0.3	0.0
31. Wind	-0.1	1.4	1.2	-0.7	-1.0	0.8
33. Laughter	2.3	1.7	1.6	-0.5	-0.6	1.1
34. Smoke	-1.7	0.5	-1.2	-1.0	-1.2	0.4
39. Truth	2.0	2.3	0.4	2.0	0.5	0.8
47. Snake	-1.2	0.8	-0.4	-0.9	-1.4	0.7

Delhi/Hindi

2. Progress	1.9	1.9	1.1	1.4	0.8	1.2
8. Anger	-1.6	0.6	0.5	-1.9	0.9	0.7
10. Crime	-2.3	0.8	-0.1	-1.6	1.0	0.4
22. Freedom	2.4	2.3	0.4	1.1	-0.3	1.1
24. Success	2.4	1.9	0.5	2.1	0.2	1.4
28. Defeat	-1.7	0.0	-.05	-1.9	-0.1	-1.1
29. Cloud	0.9	1.2	-0.4	0.0	-1.3	0.3
30. Policeman	0.9	1.2	0.4	0.5	0.3	0.0
31. Wind	-0.1	1.4	1.2	-0.7	-1.0	0.8
33. Laughter	2.3	1.7	1.6	-0.5	-0.6	1.1
34. Smoke	-1.7	0.5	-1.2	-1.0	-1.2	0.4
39. Truth	2.0	2.3	0.4	2.0	0.5	0.8
47. Snake	-1.2	0.8	-0.4	-0.9	-1.4	0.7

Finland/Finnish

2. Progress	2.0	0.7	0.6	1.0	0.7	1.2
8. Anger	-2.1	1.6	-0.7	-1.5	1.1	0.4
10. Crime	-2.2	1.6	-0.3	-1.5	1.0	0.3
22. Freedom	2.6	-0.1	1.9	1.6	-0.9	1.5
24. Success	2.2	0.2	1.1	1.8	-0.1	1.1
28. Defeat	-2.6	1.5	-2.0	-2.3	0.4	-0.7
29. Cloud	-0.5	-0.4	0.6	0.3	-1.7	-0.1
30. Policeman	-0.1	1.8	-0.1	0.4	1.1	0.2
31. Wind	0.5	0.1	1.9	-0.3	-0.7	0.8
33. Laughter	1.8	0.2	1.6	-0.1	-1.1	0.7
34. Smoke	-1.0	-0.6	1.0	-0.7	-1.0	0.6
39. Truth	1.2	1.2	-0.1	1.5	0.8	0.6
47. Snake	-1.7	0.2	1.8	-1.3	-1.4	0.3

Germany/German

2. Progress	1.9	1.6	1.5	0.7	1.3	1.3
8. Anger	-2.4	1.4	1.5	1.5	-1.8	1.3
10. Crime	-2.6	1.9	1.4	-1.9	0.4	0.3
22. Freedom	2.3	1.5	0.3	0.5	-0.5	0.8
24. Success	2.1	1.3	0.3	1.7	0.7	1.1
28. Defeat	-2.4	0.6	0.4	-2.1	-0.3	-0.8
29. Cloud	1.2	0.0	-1.5	0.4	-1.9	0.3
30. Policeman	-0.6	1.2	0.3	0.8	1.6	0.1
31. Wind	1.2	0.4	1.6	-0.7	-1.4	0.9
33. Laughter	2.0	0.4	1.6	0.3	-1.6	0.8
34. Smoke	-0.1	-0.6	-0.9	-0.6	.1.4	1.0
39. Truth	1.6	1.4	-1.1	1.7	0.3	0.0
47. Snake	-1.6	0.4	1.2	0.0	-1.7	0.8

Japan/Japanese

2. Progress	1.9	1.4	0.1	1.4	0.6	1.5
8. Anger	-1.2	1.1	0.3	-1.6	1.4	0.7
10. Crime	-2.2	0.5	0.1	-1.6	1.1	0.3
22. Freedom	2.2	1.0	0.7	0.6	-1.3	1.3
24. Success	2.4	1.6	0.5	1.9	0.2	0.9
28. Defeat	-2.4	0.1	-1.3	-2.2	-0.1	-0.3
29. Cloud	0.8	0.0	0.0	0.2	-1.7	0.7
30. Policeman	0.1	1.1	-0.7	0.3	1.2	0.4
31. Wind	0.9	0.6	1.0	-0.5	-0.9	1.2
33. Laughter	1.6	0.5	1.8	-0.5	-1.4	1.3
34. Smoke	-1.3	-0.2	0.2	-0.5	-1.2	1.2
39. Truth	1.6	1.5	-1.0	1.4	0.8	0.2
47. Snake	-1.5	0.1	0.3	-0.4	-1.3	0.2

Cross-cultural factor analysis

The 64 GD scales and the 12 SD scales for the five sites combined generate a triangular matrix of 380 correlations (76 x 5). Each correlation is based on an N of 50 (concepts) after the individual subject data are collapsed over the group of 20-25 subjects (For additional details see Osgood, Suci and Tannenbaum.[1] Osgood.[4] Miron and Osgood.[8] James.[10]) The first three factors of the principal axes solution for the 380-correlation matrix account for 54.8% of the total variance (36.3%, 11.3%, and 7.2%, respectively). Two additional smaller factors (4.2%, and 2.7%) were also extracted, but are not presented here. The percentage of total variance accounted for by the first three factors, as well as their respective magnitude, are comparable to the values obtained previously with the verbal scales of the pan-cultural SD.[10]

Table 3 groups the scales by language/culture and lists, for each Or the three factors, the highest loading GD scales above the dotted line; immediately below the dotted line are given the four SD scales previously identified as representatives of the E factor (I), the P factor (II), and the A factor (III). The numbers appearing under the "Scale I.D." column identify the particular scales: 1-64 for the GD (refer to Figure 1) and 65-76 for the SD (refer to Table 2). The positive and negative values of the loadings reflect the arbitrary (random) left-right orientation of the scales in the test booklets: a positive value goes with the left-hand side of both the SD and GD scales (as given in Figure 1 & Table 2) while a negative value connects the right hand side of the GD scale with the left hand side of the SD scale (or vice versa) as given in Figure 1 & Table 2. Factor I ("I Scales") is equally defined by both the GD scales listed and the E scales of the SD in each of the five locations. The loadings range from the low 70's to the low 90's with the Delhi-Hindi scales contributing slightly less to the definition of the factor. The high loadings of all 20 verbal E scales Or the SD and the purity of the GD scales allow for a clear interpretation of this factor as a cross-cultural Evaluation factor for combined GD and SD scales.

Factor II ("II Scales") is much less clearly defined, the scales that load highest on it are impure, and it is defined unequally by the various cultures. For three of the cultures-- Finish, German, Japanese--both GD and the SD P scales have sufficiently high pure loadings to allow for the factor's interpretation as a cross-cultural Potency factor for combined GD and SD scales, For American English, only a few GD scales contribute to the definition of this factor, while the SD P scales contribute almost nothing to it. For Hindi, one of the SD P scales has a substantial loading, but no other scale, either SD or CD, appears to contribute substantially to this factor.

The pattern for Factor III ("III Scales") is similar to that of Factor II. For Finnish, German, and Japanese, the loadings of most of the SD A scales and several of the GD scales are sufficiently high and relatively pure to allow for the interpretation of the factor as a cross-cultural Activity factor for combined GD and 53 scales. For English, two of the SD A scales and several of the GD scales make substantial contributions to this factor's definition. For Hindi, neither the GD scales nor the SD A scales appear to contribute substantially to this third factor.

The pattern of results considered so far suggests that in our attempt to find a graphic equivalent for the verbal 53 we will have greater success with the E factor than with the P and A factors. In connection with the latter, there appear to be cultural differences in the affect assigned to the GD scales that contribute to the definition of factors II and III. In particular, American-English tends to assign part of the variance of the verbal P scales to factor I, and a lesser part to factor III. Delhi-Hindi splits the variance contribution of both its P and A scales between the first and second factors. Finland-Finnish consistently assigns most of the variance of the verbal P scales to factor II, but distributes the variance of its A scales unequally over all three factors. Germany-German and Japan-Japanese exhibit more consistent patterns congruent with a P and A interpretation of factors II and III respectively, but not in every case.

The next step in the analysis consisted of selecting four particular GD scales for each of the three factors. The following criteria influenced these selections: each scale should be as "pure" as possible with high loadings on the factor it represents and minimum loadings on the other two factors; it should maintain this pattern for all five cultures; within these two restrictions, it should have higher loadings for the factor it represents than any of the other remaining scales. The four cross-cultural GD scales selected to represent the E factor correlate quite well with the E scales of the SD in every location.

The cross-cultural GD scales selected to represent the P and A factors correlate quite inconsistently with the SD scales; the majority of the values of the coefficients are not significant (with N = 50, p <.05 =.28)

The inter correlations between the composite scores for the 50 concepts computed by taking means over the four scales representing each factor using the SD and GD scales. The left hand side of the table under the column, "(N = 50)" lists the correlation coefficients. The diagonal entries provide the estimate of the equivalence better. the two instruments. The pattern previously noted in connection with the factor analysis results is obvious once again: the E ratings for concepts as indexed by the cross-cultural GD are good estimates of the E ratings given by means of the verbal SD scales; the P and A ratings for the verbal SD are not well estimated by the selected GD scales, and the degree of correspondence varies from culture to culture: the case for Finnish is better than for German and Japanese, and these in turn are more consistent than for English and Hindi.

The composite means for 13 of the 50 concepts on the three factors computed on both instruments vary from "3.0 to "-3.0". For example, SUCCESS should be rated as highly positive on the E factor, DEFEAT as highly negative. Similarly, TRUTH, FREEDOM, POLICEMAN should receive high P scores, DEFEAT, SMOKE should be low on P. WIND, LAUGHTER, PROGRESS should be high on A, while DEFEAT, SNAKE, CLOUD should be low on A. In general, the SD ratings corroborate these expectations for all five cultures. Looking now at the GD ratings, one finds a surprising degree of face validity, more so than the size of the correlations would have led one to believe. Thus, SUCCESS, PROGRESS, FREEDOM, TRUTH receive high positive scores on factor I, and high negative scores obtain for ANGER CRIME DEFEAT Similarly, DEFEAT, CLOUD, SMOKE receive low scores on factor II, POLICEMAN gets high scores on factor II, WIND, PROGRESS, LAUGHTER, are relatively high on the third factor while DEFEAT is high negative on this factor.

In general, the polarity of ratings on the GD scales is lower than on the SD scales suggesting that its "affective space" tends to be more "packed" around the origin, and therefore, less differentiated. This is particularly true of the second and third factors. There are, however, certain interesting exceptions, where the GD mean is more extreme. Inspection of Table 4 reveals that for the concepts CLOUD, SMOKE and SNAKE the CD mean on factor II is not only more extreme than the corresponding mean on the SD but is also opposite in sign in all but two cases. In every case, the means for these three concepts have a minus sign on GD factor II. Inspection of the four GD scales involved (see no. l, 15, 17, 46 in Figure l) reveals that each of the four graphic pairs is characterized by a "roundness-angularity" opposition (with the "-" sign, being aligned with the "roundness" alternative). Apparently, the "roundness" feature of the pictograph influences the ratings of the concepts CLOUD, SMOKE, SNAKE, which of course refer to things in the world that are round. In other words, we find evidence here for what has previously been called "denotative contamination" in connection with the SD (e.g. see Osgood.[9]; Miron and Osgood.[8]). For example, the scale hot-cold is used affectively (metaphorically) with the concept JAZZ, but non-affectively (literally, denotatively) with the concept SUN or with FIRE. Similarly with hard-soft for the concepts LOVE or RESEARCH versus DIAMOND or PEACH.

In order to examine the extent of visual denotative contamination of the GD scales relative to the SD, I divided the 50 concepts into two sets, one "concrete" and the other "abstract." I reasoned that the abstract concepts would be less liable to denotative contamination on the GD scales, hence their means should be closer approximations to the corresponding SD means than what would obtain with the concrete concepts.

The comparison very clearly supports the expectation: in almost every case the coefficients for the 23 abstract concepts are higher than those for the 23 concrete concepts (median absolute r:.53 vs..27). Looking at the diagonals, which is a more directly relevant comparison for the hypothesis, the same kind of difference emerges (median absolute r:.76 vs..48). There are some dramatic differences in the relation between the A factor of the SD and GD factor III:.59 vs. -.18 for AE,.78 -vs..64 for FF,.59 vs..17 for GG, and.67 vs..08 for JP. Similar, if less dramatic, differences obtain for the E and P factors. There are, however, some reversals where better correspondence is exhibited by the concrete concepts:.52 vs..88 for the E factor with AE and.35 vs..26 for the P factor with JP. There is also a curious inversion for the P factor with AE where the abstract concepts exhibit a slight negative correlation (-.20) between the two instruments. It appears, then, that although there is a strong suggestion for the presence of denotative contamination with the GD scales, the present data are not sufficiently consistent to be quite conclusive.

There is a final question that needs to be examined before going on to the discussion of the results of this study. We have noted the fact that there seem to be variations between the cultures in terms of the adequacy of the selected GD scales as estimates of the SD. The question poses itself whether the equivalence adequacy between the two instruments can be improved by selecting GD scales separately and independently for each culture. Four such "indigenous" GD scales were selected for each location by inspecting the correlation matrix for the data in each language/culture and identifying the GD scales which correlated most highly and purely with the four SD scales that represented the three factors. This search yielded four SD scales for each factor in the five cultures. As can be seen, the best indigenous scales overlap in only a small number of cases with the cross-cultural GD scales (see Table 4). Furthermore, the same scale may tap different factors in different cultures (e.g. scale 1 is A for DH but P for GG and JP).

Looking at the differences in intercorrelations between the two instruments for the abstract versus the concrete concepts, the previously noted evidence for visual denotative contamination is again corroborated here, even more strongly than before: in 12 of the 15 diagonal comparisons, the coefficients for the abstract concepts are larger (in several cases, quite substantially).

It should be noted that the procedures used in this study are likely to underestimate the degree of equivalence of the two instruments. It will be recalled that there were five sub groups of subjects, so that each subject rated only 10 concepts. In addition, the ratings were spread over two different sessions. It would be expected that under more suitable conditions (e.g. every subject rating every concept on the selected and reduced number of GD scales), the intercorrelations between the ratings with the two instruments would yield even higher coefficients than those obtained.

Discussion

The results of this study are quite encouraging. At the very least, it demonstrates that a "graphic differential" consisting of "pictographic opposites" is a possible instrument: subjects can rate concepts on such scales with readiness and ease, and they do so, as the size of many loadings indicate, with reliability. But these data go much further as regards the development of such an instrument. This study succeeded in conclusively identifying several graphic scales that index the evaluative component of concepts. Intercorrelations of composite scores between the GD and SD evaluation (E) scales run for the most part in the 80's and 90's under less than ideal conditions for

estimating equivalence. This is true for both indigenous and cross-cultural analyses.

The implications of this finding appear to me quite important and should not be underestimated. In the first place, the evaluative component of the semantic differential technique is by far the most important of the other affective components of words and is an excellent index of general attitude.[9] Thus, we now have at our disposal a cross-cultural graphic instrument for measuring attitudes, viz. a culture-free instrument which surmounts the usual difficulties involved in measuring attitude with verbal scales.

In the second place, these findings corroborate the universality of the Gestalt laws of symmetry. Inspection of the graphic scales which correlate highly with the evaluation scales (Nos. 4, 10, 25, 31, 45, etc.; see Figure 1) indicates that, in every case, the pictographic alternatives involve a contrast between a symmetrical and a non-symmetrical figure. The symmetrical side is always good, pleasant, sweet, beautiful, happy, etc. while the non-symmetrical side is bad, unpleasant, sour, ugly, unhappy, etc. This is not to say that in art symmetry is always preferred to asymmetry, but only that, under conditions where a subject is confronted with a contrast such as this on a rating scale, he uses such a scale according to the laws of good Gestalt by apparently abstracting the common evaluative element between the concept that is rated and the symmetrical contrast of the pictographs. That untrained and naive subjects in five quite different cultures should exhibit this tendency with such high agreement is indeed a remarkable fact.

On the other hand, our success in identifying graphic scales that tap the Potency and Activity components Or affect is much less impressive. There may be two reasons for this. One is that these affective components are less stable cross-culturally, even when verbal scales are used. The other reason relates to the specific choice of pictographs included in this study. The "physiognomic quality" of these graphic scales is simple and one-dimensional, and they are all black on white. We have evidence.[10] that some dimensions of color are good correlates of Potency and Activity (saturation and brightness, respectively) and, before we give up the attempt, these additional stimulus dimensions should be tried. It may be, that visual form per se relates to evaluation only, and additional variations (three-dimensions, color) are needed to tab the other affects and to reduce denotative contamination. The question remains unsolved.

Nevertheless, the limited success obtained of indexing Potency and Activity in some cultures should not be overlooked. Thus, with selected indigenous scales, and with abstract concepts, the equivalence coefficients between the two instruments also runs in the 80's for AE, FF, and JP (for Potency) and for FF and GG (for Activity). These results, however, because they are not uniform, are to be taken with caution and need replication.

Previous research with the semantic differential technique has established the cross-cultural generality of three basic affective dimensions identified as Evaluation, Potency, and Activity. The purpose of this study is to develop a graphic equivalent of the pan-cultural short form semantic differential. This "graphic differential" is composed of bi-polar scales made up of pictographic opposites on which subjects rate a concept by placing a check mark on the seven-point scale. Such a non-verbal instrument would be culture free and would, under several conditions, have certain distinct advantages over the verbal semantic differential form. Data are reported for- a pre-test of the graphic instrument in five language/culture communities (American English, Delhi Hindi, Finnish, German, and Japanese).

The study succeeds in identifying several graphic scales that provide reliable and effective indices of the Evaluation dimension of affect. These scales exhibit.' and conform to the symmetry law of Gestalt and could be used as a culture free attitude measure. With respect to the Potency; and Activity dimensions, the results are less satisfactory, although some of the graphic scales can be used for this purpose when chosen indigenously (i.e. independently for each culture). Evidence of visual denotative contamination is found and discussed.

Acknowledgments

Conflicts of interest

The authors declare that there is no conflict of interest.

References

1. Osgood CE, Suci G, Tennenbaum PH. The measurement of meaning. University of Illinois Press, Urbana, USA; 1957. p. 1−360.

2. Snider JG, Osgood CE Semantic Differential Technique. Aldine Press, Chicago, USA; 1969.

3. James LA. Studies in comparative psycholinguistics. *International Journal of Psychology.* 1966;1:15−37.

4. Osgood CE. Semantic differential technique in the comparative study of cultures. *American Anthropologist.* 1964;66(50):171−200.

5. Center for comparative psycholinguistics a cross-cultural atlas of affective meanings. Institute of Communications Research, the University of Illinois, Urbana, USA; 1959.

6. Osgood CE. The cross-cultural generality of visual-verbal synesthetic tendencies. *Behavioral Science.* 1960;5(2):146−169.

7. Brown roger social psychology. The-Free Press, New York, USA 1965.

8. Miron MS, Osgood CE. The multivariate structure of qualification. In: Cattell RB, editor. Handbook of Multivariate Experimental Psychology. Rand McNally, Chicago, USA; 1950.

9. Osgood CE. Studies on the generality of affective meaning systems. *American Psychologist.* 1962;17(1):10−28.

10. James LA. Some potential uses of the cross-cultural atlas of affective meanings. center for comparative psycholinguistics, The University of Illinois, Urbana, USA; 1967.