INTERNATIONAL CONFERENCE ON ENGINEERING DESIGN, ICED'07
28 - 31 AUGUST 2007, CITE DES SCIENCES ET DE L'INDUSTRIE, PARIS, FRANCE

CAN AN OBJECTIVE MEASUREMENT OF DESIGN PROTOCOLS REFLECT THE QUALITY OF A DESIGN OUTCOME?

Jeff WT Kan[1] and John S Gero[2]

[1] Key Centre of Design Computing and Cognition, University of Sydney, Australia
[2] Krasnow Institute for Advanced Study and Volgenau School of Information Technology and Engineering, George Mason University, USA

ABSTRACT

In this paper we develop a method for measuring the design process and correlate it with the assessment of the design outcome. We used linkography – a technique used in analyzing design protocols – which is reported to reveal the quality of design ideas and to capture the interconnectivity of ideas in the design process. Previous studies found that Shannon's entropy as a measure of a linkograph might indicate design outcomes. This paper reports on an empirical verification that design outcomes can be correlated to the entropic interpretations of a linkograph. Experimental data from protocol studies of six designers performing two similar design tasks in two different conditions were analysed. In one condition the designers were blindfolded during designing, and then given a relatively short time to draw their design at the end of the session. In the other condition the designers were allowed to sketch. We focus on the three highest scored and the three lowest scored design sessions. Their verbalization during the design session were segmented and coded to produce linkographs. The entropy and the change of entropy over time of these linkographs were studied in relation to the design outcomes. Our results demonstrate that entropy measurement of linkographs provides a way of to study design protocols and show that the design outcome is more related to the rate of change of entropy than to the entropy alone.

Keywords: design measurement; entropy; design protocols; design outcomes

1 INTRODUCTION

There is interest in assessing the design product in association with the design process, with studies trying to find the relationship between design performance and integration of ideas, learning style, and design strategies [1-3]. In the process of designing, freehand sketching has traditionally been seen as the core activity for designing despite the development of digital tools [4, 5]. However, results from recent studies have challenged the role of sketching in conceptualization and designing [6, 7]. In order to study the effects of not being able to sketch during designing, we develop a method for measuring the design process and correlate it with the assessment of the design outcome. We explore the use of linkography, a technique used in protocol analysis of designing, to capture the structure of design idea generation processes. We then measure the linkographs, based on information theory, and correlate this with the design outcome. The goal is to test, empirically, whether an entropic measurement of linkographs can be used as a quantitative, objective, tool to access the "goodness" of design processes, ideas, and outcomes.

2 METHOD

This section briefly explains how two sets of design protocols from two different conditions were collected, the criteria of judging design outcomes, the method of coding linkogaphs, and how to use entropy to measure the linkographs. We hypothesize that entropy measurement of the design sessions can indicate the success of design outcomes.

2.1 Experiments

We are using the data from Bilda's [8] study. In that study, six architects each with over 10 years of experience were involved in two design sessions of about 45 minutes each, with one month separation between them. Both sessions involved designing a house in the same site with different requirements. In both sessions the architects were ask to talk-aloud while designing. In one of the sessions the designer was blindfolded (BF) during designing so that s/he was unable to sketch, and then s/he was asked to quickly draw what was designed at the end of the session with the blindfold taken off. In the other session, the same architect designed normally, that is, using sketching (SK). The sessions were video and audio recorded as raw protocol data. In order to be objective in the assessment, the design outcomes of the six participants, 12 design sessions, were double-blind reviewed by three judges according to five categories: innovation, creativity, satisfaction of design brief, practicality, and flexibility. All the judges have practiced and taught design for more than 15 years [8]. The judges were unaware of the experiment.

2.2 Constructing the linkograph

Linkography was introduce by Goldschmidt [9] and has been used for investigating the structure of design idea generation processes and for comparing design productivity [10, 11]. It is a graphical representation of design processes in terms of the links between design moves. A move, as defined by Goldschmidt, is a step, an act, an operation, which transforms the design situation relative to the state in which it was prior to that move. The protocols were segmented into moves and each move was discerned if it was to be linked to previous moves by the coder. Figure 1 is an example of a linkograph with four moves used by Goldschmidt [1].

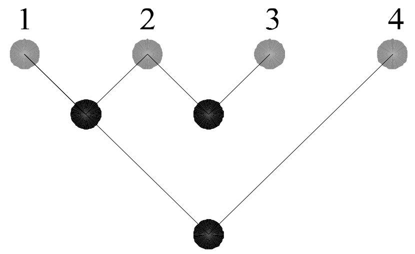


Figure 1. Linkograph of the four moves with moves as grey dots and black dots denoting links. Move 1 on the left is linked to moves 2 and 4. Move 2 is linked to move 3.

To illustrate the construction of a linkograph the following three moves were extracted from the protocol of a session in this study.

51. (15.50) Look, the thing that I'm thinking now is that because I've got such an overwhelming desire to design a courtyard house, and I think that in this kind of situation where you've got a very large site and, umm, a semi-public space that it can borrow, in a way, (16.07) that what you'd start to plumb for is a courtyard building; parts of which are built and parts of which are unbuilt.
52. (16.14) So, I'd be inclined to organise the dancer's studio and the living spaces and parts of the, the bedrooms…(16.29) or no the bedroom spaces I think should go down to the east…to give them some separation…
53. (16.34) So I'm imagining now a broken form, something that's got the courtyard essentially as its organising structure, but which then has parts built, parts unbuilt.

In move 53, the architect refers and develops the courtyard house idea in move 51 therefore a link is established between move 51 and 53.

The pattern, distribution, and intensity of links can be used to assess the design process. Goldschmidt identified two types of links: backlinks and forelinks. Backlinks are links of moves that connect to previous moves and forelinks are links of moves that connect to subsequent moves. Conceptually they are very different: "backlinks record the path that led to a move's generation, while forelinks bear evidence to its contribution to the production of further moves" [10].

2.3 Entropy measurement of linkograph

Kan and Gero [12-14] used Shannon's entropy [15] to quantify a linkograph on the basis that fully linked and empty linked linkographs represent substandard design processes. An empty linked linkograph represents neither development nor consolidation of ideas; a fully linked linkograph suggests fixation and no diversification of ideas. In Shannon's information theory, the amount of information carried by a message or symbol is based on the probability of its outcome. If there is only one possible outcome then there is no additional information because the outcome is known. Information, carried by a symbol, can then be defined related to the surprise the symbol produces or the decrease in uncertainty by the symbol. For an *n* symbols system, the average information per symbol, entropy *H*, is derived by Shannon as formula (1), where p_i is the probability or the frequency of occurrences of symbol *i*.

$$H = -\sum_{i=1}^{n} p_i log\,(p_i) \quad with \quad \sum_{i=1}^{n} p_i = 1 \qquad (1)$$

Here we consider two symbols: links as ON and not linked as OFF, this is sufficient to describe the states of a move in relation to other moves in a linkograph. We calculate the entropy of a linkograph by using formula (2) and apply it to each move. Based on the conceptual difference of forelinks and backlinks we measure them separately.

$$H = -p(ON)Log(p(ON)) - p(OFF)Log(p(OFF)) \qquad (2)$$

H will be zero if *p(ON)* equals 1 or *p(OFF)* equals 1. *H* will have a highest value of 1 when *p(ON)* equals *p(OFF)* equals 0.5.

We introduce another link type, horizontal link (called horizonlink), which carries the notion of cohesiveness between linked moves. Each grey rectangle in Figure 2 is a basic entity in our entropic calculation. In the case of the forelink and backlink, each entity represents the contribution of an individual move. We add the entropy contributed by these entities to obtain the forelink (FH), backlink (BH), and horizonlink (HH) entropy of a linkograph.

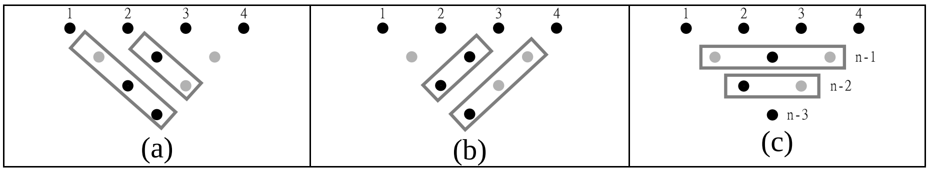


Figure 2. Abstracted linkograph for entropy measurement, black dots denote links. (a) Measuring entropy of forelinks, (b) measuring entropy of backlinks, and (c) measuring entropy of horizonlinks.

If the idea in a move is weak, it will not have a lot of forelinks (*p(OFF)* is close to 1) and this is produces a low entropy. However, if an idea gets too many forelinks (*p(ON)* is close to 1), this might indicate fixation; which is also indicated by a low entropy. Similarly, backlink entropy measures the opportunities according to enhancements or responses. If an idea is very novel, it will not have backlinks (*p(OFF)* equals 1), the resulting entropy is zero. If an idea is backlinked to all previous ideas (*p(ON)* equals 1), it lacks novelty and is represented by zero entropy. Moves that are close will usually have better interconnectivity; we refer this as the cohesiveness of moves. Links that connect moves that are far apart, we consider those as incubated linked moves. The measurement of horizonlink entropy will encourage the occurrence of incubated segments and discourage too strong cohesiveness.

Our hypothesis is that the entropy measurement of a linkograph is positively correlated to the design outcome.

3 RESULTS

3.1 Design Outcome

Table 1 summarizes the averaged score of the assessment criteria of the 12 design sessions, where BK and SK refer to blindfolded and sketching respectively and the number following is the designator of the particular designer. We can observe that: 1) the average score of the BF sessions is higher than the SK sessions; 2) all the BF sessions received higher scores than the corresponding SK sessions with one exception that had an equal score; 3) the scores range from 3.7 to 7.7 with the average 5.87 and standard derivation of 1.12. There are two 3.7 scores which is the lowest and three 4s. For the highest there are one 8 and four 7.7s clustered around the blindfold session in the practical and satisfying the design brief category. The two 3.7s and three 4s scatter in different categories.

TABLE 1. Cumulative score of the criteria in design sessions by three judges

	Innovative	Creative	DesBrief	Practical	Flexible	Average
BF01	4.0	5.3	7.7	7.7	6.0	6.1
BF02	4.3	6.0	6.3	7.0	6.3	6.0
BF03	6.0	6.3	7.7	7.0	7.3	6.9
BF04	5.0	5.7	7.5	6.7	5.7	6.1
BF05	6.3	7.3	8.0	7.7	6.0	7.1
BF06	4.3	3.7	5.7	5.0	5.7	4.9
SK01	4.3	5.0	6.3	6.0	5.3	5.4
SK02	5.3	5.7	6.3	5.7	6.3	5.9
SK03	6.7	7.3	6.3	5.3	6.7	6.5
SK04	4.3	4.7	5.0	3.7	4.0	4.3
SK05	6.0	6.3	7.0	7.0	5.7	6.4
SK06	4.0	4.7	5.3	5.7	5.0	4.9

3.2 Entropy Measurement of the Design Sessions

Here, only the three highest scored and the three lowest scored sessions will be studied in more detail, the sessions are BF05, BF03, SK03 and SK04, SK06, BF06 respectively. The score difference between these two groups is over 40% which gives us a good ground for comparison. The innovative and creative categories are the main contribution to this score difference. Results of the other six middle scored sessions will be presented but not considered in detail because the difference among them is not significant. Table 2 shows the normalized entropy of each session together with their overall score. For the ease of comparison we use the normalized value: linkograph entropy divided by the number of moves, instead of absolute values.

Overall the BF sessions have higher entropy than their corresponding SK session with one exception. The differences in entropy are marginal and the evidence is insufficient to suggest correlation and conclude that BF sessions have higher entropy than their corresponding SK session. Positive correlation between entropy and the evaluation of the design outcome is weak. Some of the lowly ranked design sessions have high entropy values.

TABLE 2. Entropy and link index of each session

	nBH	nFH	nHH	Total	Outcome
BF01	0.125	0.122	0.060	0.307	6.1
BF02	0.161	0.155	0.066	0.383	6.0
BF03	0.143	0.140	0.055	0.338	6.9
BF04	0.240	0.220	0.093	0.553	6.1
BF05	0.224	0.193	0.082	0.499	7.1
BF06	0.188	0.189	0.105	0.481	4.9
SK01	0.137	0.124	0.077	0.337	5.4
SK02	0.157	0.150	0.065	0.373	5.9
SK03	0.124	0.131	0.044	0.299	6.5
SK04	0.227	0.203	0.098	0.529	4.3
SK05	0.176	0.125	0.071	0.372	6.4
SK06	0.184	0.175	0.063	0.422	4.9

nBH: normalized backlink entropy
nFH: normalized forelink entropy
nHH: normalized horizonlink entropy

3.3 Entropies of the High and Low Scored Sessions

Table 3 shows the scores and normalized total entropies of the three high scored and the three low scored outcomes of sessions in the descending order of entropy. The low scoring sessions have higher

entropies than the high scoring sessions which indicate more links among moves. This suggests that more links in the design process does not necessary produce better designs. SK04, the lowest scored session, was considered the least “practical” and least “flexible” design solution. It also scored lowest in terms of fulfilling the design brief. BF06 was judged as the least “creative” design solution.
At the other end, the high scoring sessions, BF05 and SK03 were considered to be the most creative. BF05 scored highest in terms of fulfilling the design brief, and share with one other session to be considered as the most “practical” design.

TABLE 3. Entropy and outcome, in decreasing order of entropy

High score sessions	Score	H
BF05	7.1	0.499
BF03	6.9	0.338
SK03	6.5	0.299

Low score sessions	Score	H
SK04	4.3	0.529
BF06	4.9	0.481
SK06	4.9	0.422

3.4 Change of Entropy over Moves

It is possible to infer that the entropy varies across the time line as the clusters of links are not uniform. There are two approaches to measure this change, one using a fixed time frame as a reference window and the other using a fixed number of moves as the width of window. The latter was used as there is no direct map of moves to the time scale. This will likely give more meaningful comparisons as the duration of moves are not uniform; and in terms of designing, moves are the determining factor rather than time. When calculating the entropies within a 7 moves window as in Figure 3, we can record the changes of entropy across the design session. In this method we neglect those linked nodes outside the window, which is outside the shaded triangle [16]. Kan and Gero [16] observed that with a 7 moves window it will be hard to observe the trend because too many of the links were ignored. Increasing the width of the window smooths the graph.

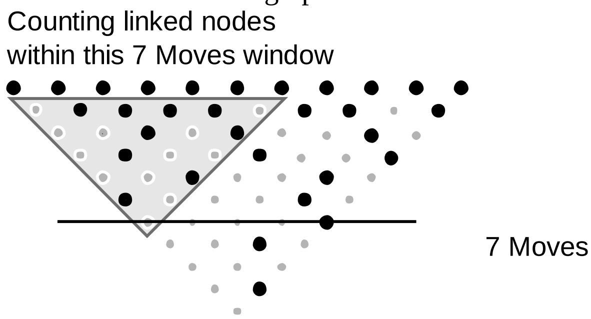


Figure 3. Using 7 moves as a window to study the change of entropy over time in a linkograph. For clarity, the links are not drawn as in normal linkographs.

Figures 4, 5, and 6 show the change of entropy over the session with a 28 moves window for the high scored sessions. The entropy drops in the middle and rise toward the end. Figures 7(a), (b), and (c) show the second degree polynomial fit of Figures 4, 5, and 6 respectively. The entropy at a given window will be highest when the links are most random, that is half linked and half unlinked. Full links or empty links will result in zero entropy; so there are two reasons for an entropy drop: saturation of links or sparse of links. Reviewing the linkographs, all fit into the second reason. So for the high scoring sessions, the beginning and the end have more links within the 28 move window.

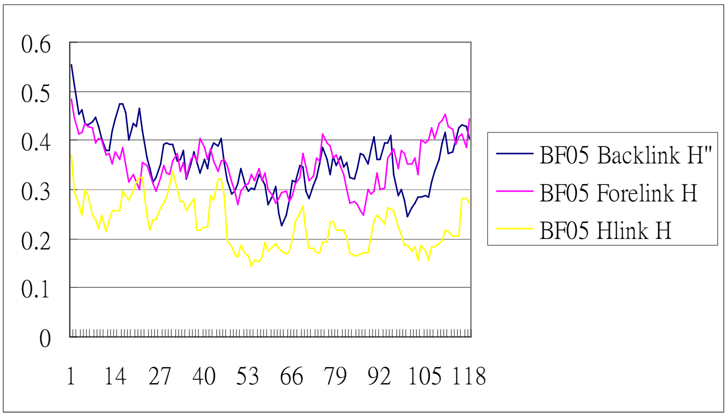


Figure 4. Change of entropy plot of design session BF05 with total score of 7.1 which is the highest

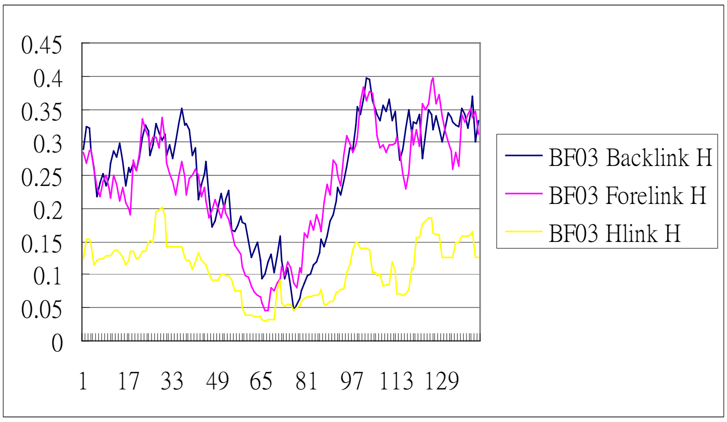


Figure 5. Change of entropy plot of design session BF03 with total score of 6.9

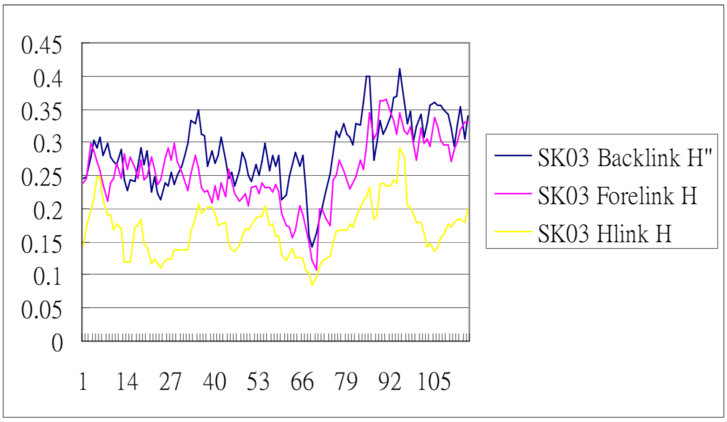


Figure 6. Change of entropy plot of design session SK03 with total score of 6.5 which has the highest innovation plus creative score

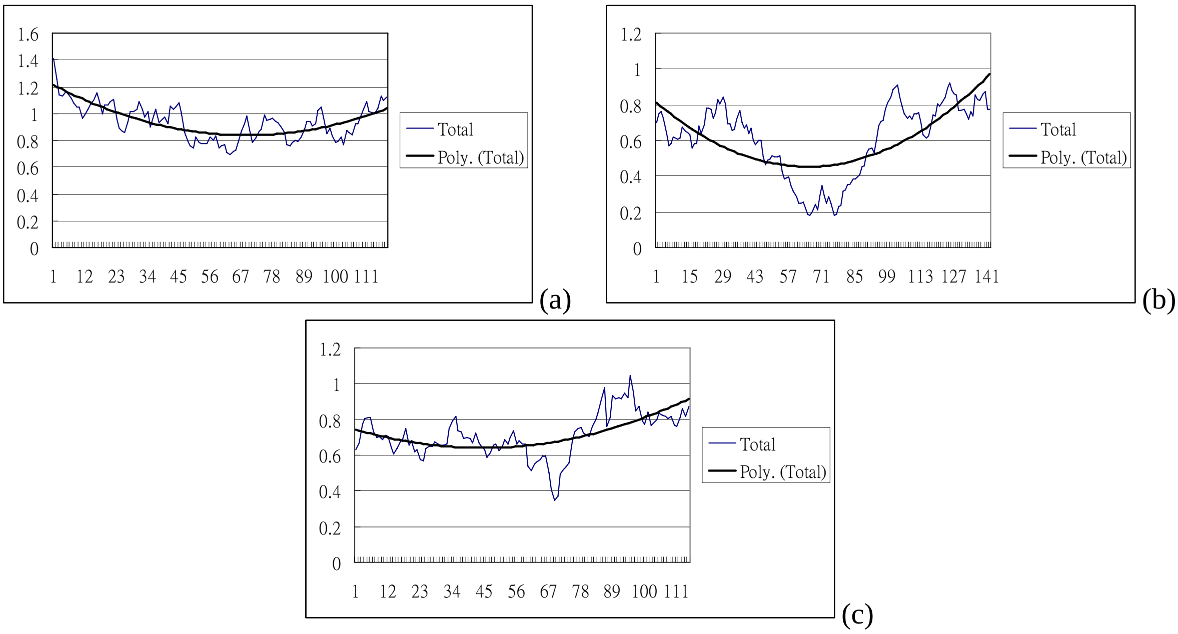


Figure 7. Quadratic fit of the change of entropy of the top three design sessions for sessions: (a) BF05; (b) BF03 and (c) SK03.

Figures 8, 9 and 10 are the plots of the change of entropy of the low scoring sessions. Figures 11(a), (b), and (c) show the second degree polynomial fit of Figures 8, 9, and 10 respectively. This is opposite to the high scoring session, at the beginning and towards the end there is less saturation of links within the 28 moves window, which means less coherences of ideas at beginning and end of the session.

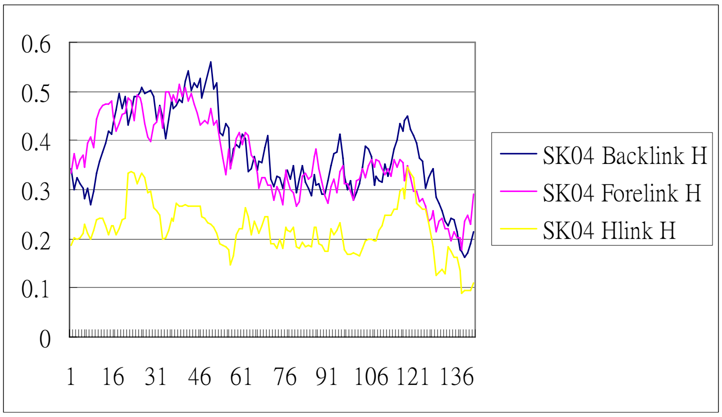


Figure 8. Change of entropy plot of design session SK04 with total score of 4.3 which is the lowest

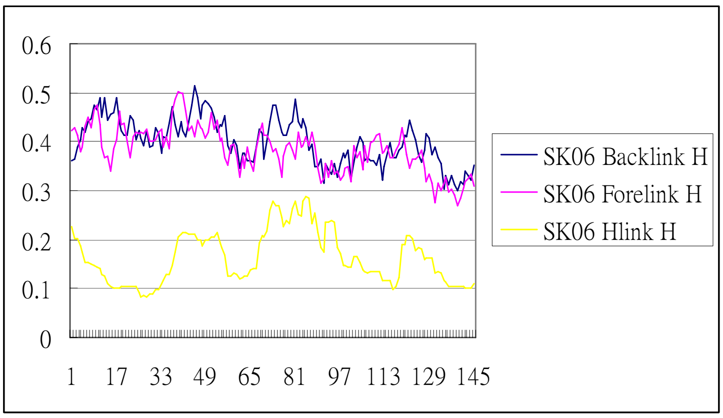


Figure 9. Change of entropy plot of design session SK06 with total score of 4.9

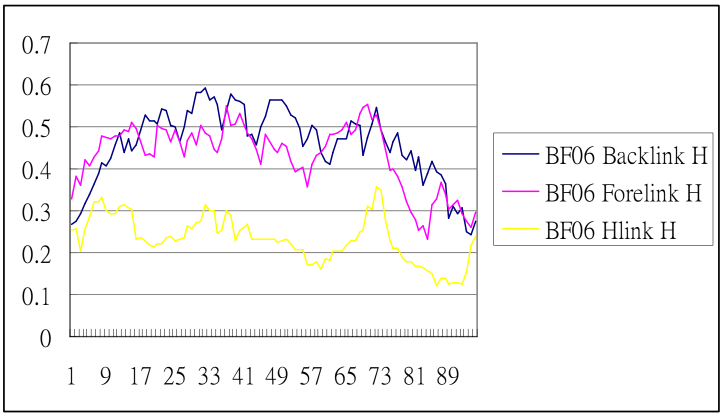


Figure 10. Change of entropy plot of design session BF06 with total score of 4.9

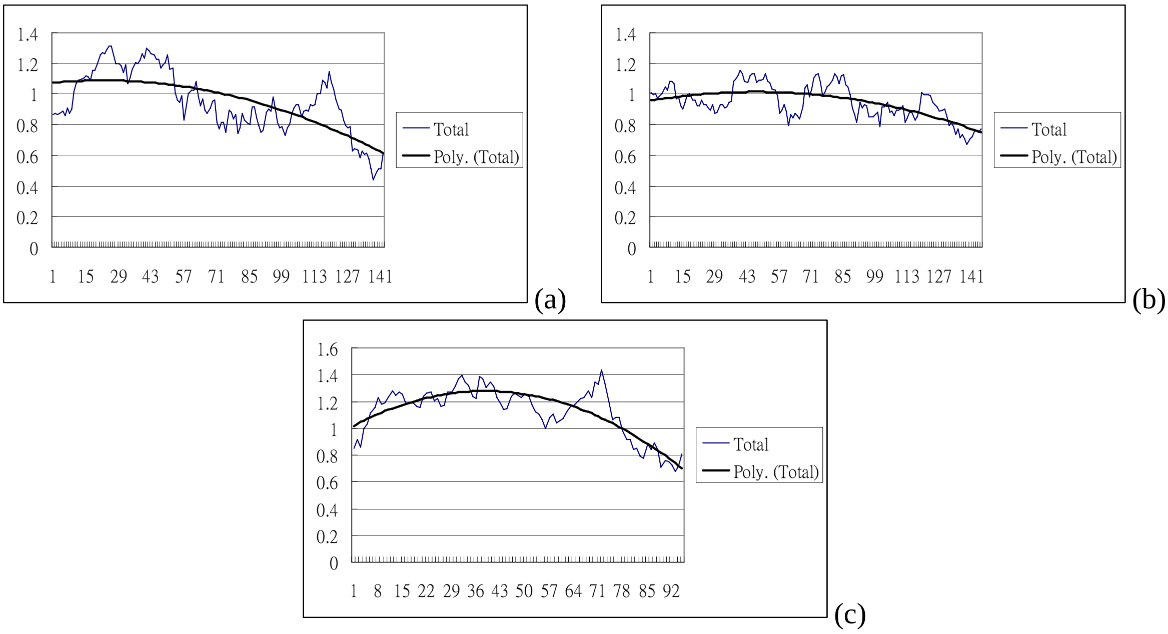
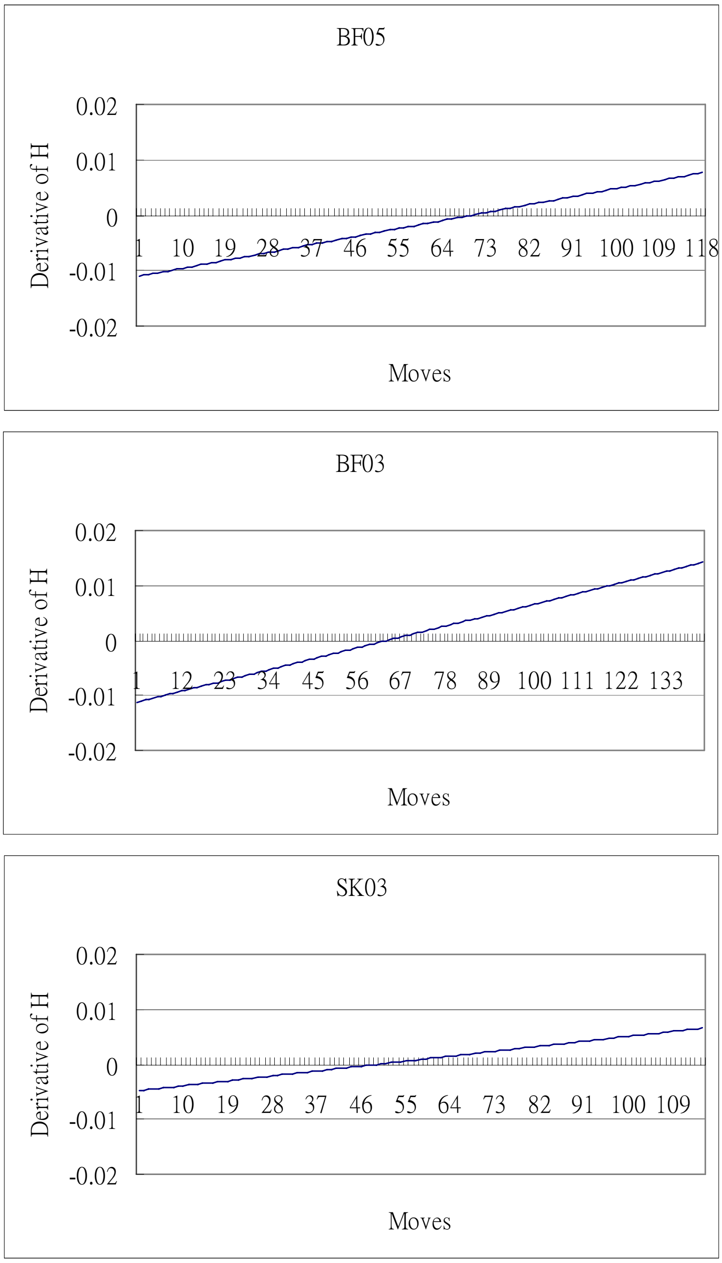


Figure 11. Quadratic fit of the change of entropy of the bottom three design sessions for sessions: (a) SK04, (b) SK06 and (c) BF06

It can be observed that all the high scoring sessions have "u" shaped or negative curvature quadratic fit curves and all the low scoring sessions have "n" shaped or positive curvature curves. If we take the differential (slope or tangent) of the straight line. The differential of the

entropy curve denotes the rate of change of entropy, Figures 12(a), (b), and (c) show the differentials of the top three sessions and Figures 13(a), (b), and (c) show the differentials of the bottom three sessions. The slopes of the differentials in Figure 12 are all positive. The rate of change of entropy increases from negative to positive and becomes zero near the middle of the session. This is the opposite in the low scoring sessions, where the rate of change of entropy decreases from positive to negative, the slopes of the differentials are all negative, Figure 13.



(a)

(b)

(c)

Figure 12. Differential of the top three sessions for sessions: (a) BF05, (b) BF03 and (c) SK03.

For the high scoring outcome design sessions the rate of change of entropy increases from negative to positive while for the low scoring design sessions the rate of change of entropy decreases from positive to negative within the forty minutes session. A positive rate of change indicates moving towards more links that is a better integration of ideas. Moving towards less links implies less integration of ideas that will result in a negative rate of change. In Figure 12, the high scoring sessions, the positive slope graphs suggest there is an increase of integration of ideas after about the middle of the session. While in Figure 13, the low scoring sessions, they are moving toward less integration of ideas toward the end of a session.

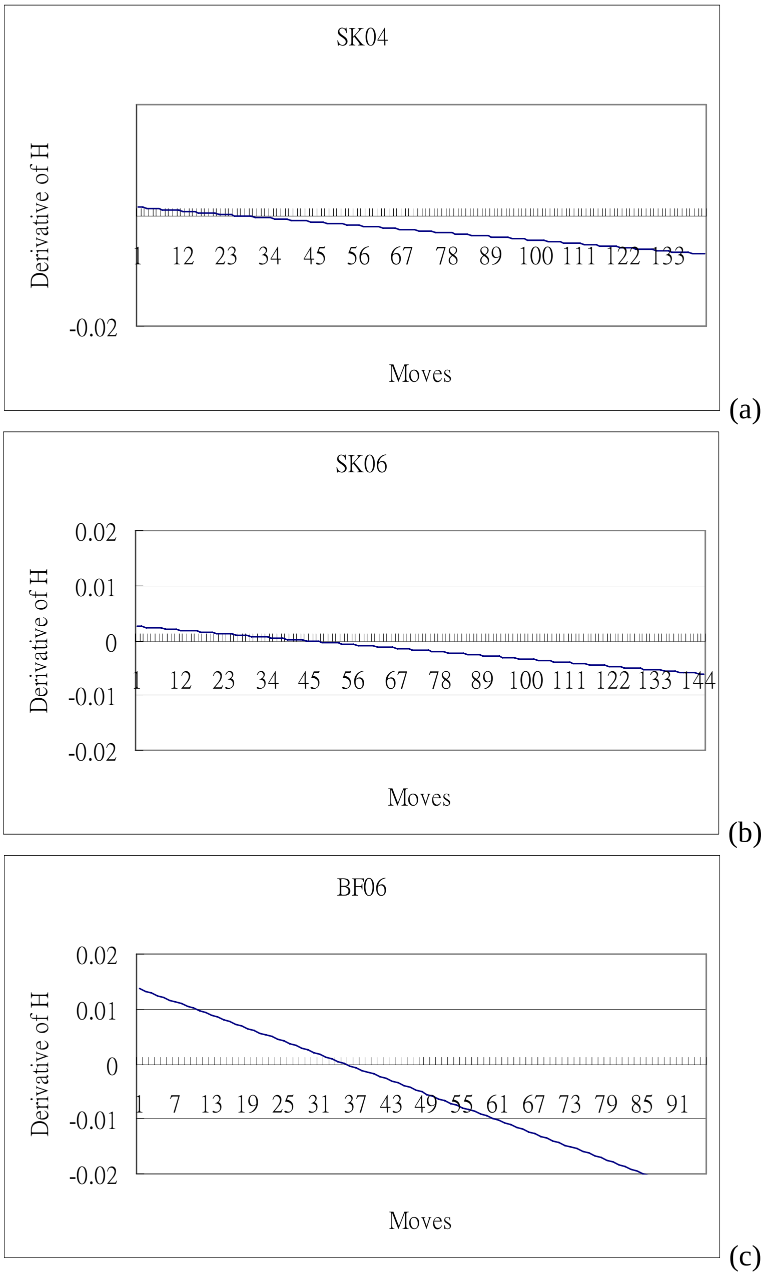


Figure 13. Differential of the bottom three sessions

3.5 Good and Bad Design Moves

Many researchers believe that there are prerequisites for the creation of useful ideas; among those experience and interactions play an important role [17-19]. A good design idea not only fulfils the requirements but also has the quality of novelty and creativity. Finke et al [19] considered creativity not as a single unitary process but a product of many types of mental processes collectively setting the stage for creative insight and discovery. If good design ideas exist, bad design ideas co-exist in relative terms. Bad ideas are those that are impractical, non-innovative, or unrealistic thoughts.

A move in a linkograph does not have any attribute nor does it have any value judgment assigned to it. We assume there are two fundamentally different types of moves, good and bad moves, where good moves contribute to good design ideas and bad moves play a part in bad design ideas. An integration of good design ideas will produce good design outcome and an integration of bad design ideas will produce a poor design outcome. If a design solution is based on poor assumptions or unrealistic desires, the outcome would not be ranked highly. In Table 1 we see the most impractical design (SK04), which also scores the lowest in outcome, has a relative high linkograph entropy, Tables 2 and

3. Could this designer have integrated his/her design based on some bad moves? We do not have the raw video data and transcript to carry out a detailed analysis, but it may be that the designer has integrated his/her design with some impractical ideas. This can be illustrated by a school project of a year one architecture student designing a house in a historic village in Hong Kong the first author was involved with. The original walled portion is a declared monument, protected from alterations. However, this student had the idea of "breaking through traditions" which included demolishing the monument walls where she integrated her design. As a result her design outcome received a lot of criticisms; one external examiner recommended failing her.

Introspectively, expert designers are able to judge the goodness of an idea and good ideas will get the designers excited while bad ideas will cause designers to lose interest. We suspect that when a designer is excited, there will be more integration in the design protocols, which results in a higher entropy. Conversely, bad design moves will reduce development potential causing entropy to drop.

Using the concept of good and bad moves, good ideas produce more integration towards the end of a session and bad ideas less integration towards the end. Does this mean that the design outcome is correlated to the rate of change of entropy? With these limited cases, it is inconclusive. However, it does indicate that there is a possible correlation. It is temping to draw the conclusion that the trend of entropy reveals: 1) the outcome of design, 2) the goodness of ideas, and 3) the design process. More experiments are needed to verify the claim that entropy can measure the richness of design outcome. These experiments can be studies of experts and novices (assuming experts will produce better design outcome than novices), and design students with high and low grades.

CONCLUSION

The aim of this exploration is to verify the entropic measurement of design processes proposed by Kan and Gero [14, 16]. The design process is captured and represented by coding design protocols into a linkograph. The linkographs are measured using Shannon's entropy. We find the design outcomes have certain relationships with the entropy of their linkograph. All the blindfolded (BF) sessions have a higher or equal score of design outcome compared to their corresponding sketch (SK) sessions. This is reflected by the entropy of the linkographs – five out of six of those BF sessions have higher entropy than the corresponding SK session. However higher entropy does not correlate with better design outcomes; it is inadequate to explain the experimental results. We introduce the concept of good and bad moves to account for this seemingly contradictory observation. All the high scoring sessions, with good moves, have higher entropic measure toward the end of the design session as they become more integrated approaching the end. This can be approximated by a quadratic fit with a negative curvature change of entropy curves. The differentials of these quadratic curves are straight lines. The three highest score sessions have a positive slope while all the poor score sessions have a negative slope.

If entropy measures idea generation opportunities, then entropy should drop towards the end of a session because as the designer approaches "a finished design" there should be fewer opportunities for ideas. The increase in entropy at the end of a session means a better integration of moves which might indicates that the moves contribute to good ideas.

Although it seems to be a jump to conclude that the entropy of linkograph is correlated with the outcome of designing as there are insufficient cases to validate the claim, there is some evidence that suggests the design outcome could be related to entropy values of linkograph and the goodness of moves which can be reflected in the pattern of the change of entropy over a session.

ACKNOWLEDGEMENTS

This research supported by an International Postgraduate Research Scholarship and a University of Sydney International Postgraduate Award. Facilities were provided by the Key Centre of Design Computing and Cognition. We would like to thank Zafer Bilda for providing the coded protocol data for generating the initial linkographs and calculating the corresponding entropies.

REFERENCES

[1] Goldschmidt, G. Criteria for design evaluation: a process-oriented paradigm. In Kalay, Y.E., ed. *Evaluating and Predicting Design Performance*, pp. 67-79 (John Wiley & Son, Inc., New York, 1992).

[2] Kvan, T. and Yunyan, J. Students' learning styles and their correlation with performance in

architectural design studio. *Design Studies*, 2004, **26**(1), 19-34.
[3] Kruger, C. and Cross, N. Solution driven versus problem driven design: strategies and outcomes. *Design Studies*, 2006, **27**(5), 527-548.
[4] Lawson, B. and Loke, S.M. Computers, words and picture. *Design Studies*, 1997, **18**(2), 171-183.
[5] Schon, D.A. *The Reflective Practitioner: How Professionals Think in Action*. (Avebury, Ashgate, London, 1991).
[6] Jonson, B. Design ideation: the conceptual sketch in the digital age. *Design Studies*, 2005, **26**(6), 613-624.
[7] Bilda, Z., Gero, J.S. and Purcell, T. To sketch or not to sketch? That is the question. *Design Studies*, 2006, **27**(5), 587-613.
[8] Bilda, Z. The Role of Externalizations in Design: Designing in Imagery Versus Sketching. *Working Paper,* Key Centre of Design Computing and Cognition (The University of Sydney, Sydney, 2006).
[9] Goldschmidt, G. Linkography: assessing design productivity. In Trappl, R., ed. *Cyberbetics and System '90*, pp. 291-298 (World Scientific, Singapore, 1990).
[10] Goldschmidt, G. The designer as a team of one. *Design Studies*, 1995, **16**(2), 189-209.
[11] Van-der-Lugt, R. Relating the quality of the idea generation process to the quality of the resulting design ideas. *International Conference on Engineering Design (ICED)*, pp. 19-21Stockholm, Sweden, 2003).
[12] Kan, J.W.T. and Gero, J.S. Can entropy indicate the richness of idea generation in team designing? *CAADRIA05*, pp. 451-457 (New Delhi, India, 2005).
[13] Kan, J.W.T. and Gero, J.S. Design behaviour measurement by quantifying linkography in protocol studies of designing. In Gero, J.S. and Lindemann, U., eds. *Human Behaviour in Designing'05*, pp. 47-58 (Key Centre of Design Computing and Cognition, University of Sydney, Australia, 2005).
[14] Kan, J.W.T. and Gero, J.S. Acquiring information from linkography in protocol studies of designing. *Design Studies*, 2007, to appear.
[15] Shannon, C.E. A mathematical theory of communication. *The Bell System Technical Journal*, 1948, **27**, 397-423.
[16] Kan, J.W.T. and Gero, J.S. Entropy measurement of linkography in protocol studies of designing. In Gero, J.S. and Bonnarde, N., eds. *Studying Designers'05*, pp. 229-245 (Key Centre of Design Computing and Cognition, University of Sydney, 2005).
[17] Cross, N. Creative Thinking by Expert Designers. *The Journal of Design Research*, 2004, **4**(2), http://jdr.tudelft.nl/
[18] Minneman, S.L. and Harrison, S.R. Negotiating right along: an extended case study of the social activity of engineering design. In Duffy, A.H.B., ed. *The Design Productivity, pp. 32-50* (Springer, London, 1998).
[19] Finke, R.A., Ward, T.B. and Smith, S.M. *Creative Cognition*. (MIT Press, Cambridge, MA, 1992).

Contact: Jeff Kan
Key Centre of Design Computing and Cognition
Faculty of Architecture, Design, and Planning
University of Sydney
Sydney, NSW 2006 Australia
Phone: +61 2 90365496
Fax: +61 2 93513031
e-mail: kwaitak@arch.usyd.edu.au