Running head: Interjudge Inconsistency

Interjudge Inconsistency Index for Body of Work, Yes/No, and Bookmark Standard Setting Procedures

Abdullah Ferdous

Measured Progress

Barbara Plake

University of Nebraska-Lincoln

Interjudge Inconsistency Index for Body of Work, Yes/No, and Bookmark Standard Setting Procedures

Abstract

A number of standard setting methods have been developed over the last 30 years (Berk, 1986; Cizek, 2001; Hambleton & Powell, 1983; Kane, 1994; Livingston & Zieky, 1982; Lunz, 1995). Some of them are more widely used than the others. For example, the Bookmark (Mitzel, Lewis, Patz, & Green, 1986) procedure is the most popular standard setting procedure used in state assessment programs (CCSSO, 2001). On the other hand, the Angoff (1971) procedure and its variations are more popular in licensure and certification programs (Sireci & Biskin, 1992). Although these procedures are very popular, limited research studies have been carried out to evaluate these procedures in terms of accuracy or consistency (e.g., intrajudge consistency and interjudge agreement or disagreement). This study developed a generalized interjudge inconsistency index that can be used for Body of Work (Kinston, Kahl, Sweeney, & Bay, 2001), Bookmark, and a Yes/No variation of Angoff (Plake, Ferdous, Impara, & Buckendahl, 2005) standard setting procedures.

Interjudge Inconsistency Index for Body of Work, Yes/No, and Bookmark Standard Setting Procedures

Many standard setting methods developed over the last 30 years have been used for setting cut scores on high-stakes test (Berk, 1986; Cizek, 2001; Hambleton & Powell, 1983; Kane, 1994; Livingston & Zieky, 1982; Lunz, 1995). Some of them are more widely used than the others. For example, the Bookmark (Mitzel, Lewis, Patz, & Green, 1986) is the most popular standard setting method used in state assessment program (CCSSO, 2001). On the other hand, the Angoff (1971) has been used in the National Assessment of Educational Progress (Reckase, 2000) and is more popular in licensure and certification programs (Sireci & Biskin, 1992). Although these methods have been used widely, limited research studies have been carried out to measure how consistently the judges make their decisions (i.e., intrajudge consistency), how much they are agreed or disagreed as a panel (i.e., interjudge consistency), and how an individual judge's ratings compare to those of the rest of the judges in the panel, for example.

Hambleton (2001) emphasized compiling validity evidence and technical documentation for the standard setting procedure: "…evidence of internal validity is needed including evidence that addresses the extent of intrapanelist and interpanelist consistency, the variability of performance standards across subpanels, and agreement between panelist's judgment…(p. 104)." Kane (1998) also stated that validity of standards should be justified not only by an independent criterion; rather, validity evidence should also be derived from the method itself (i.e., reporting inter and intrajudge consistency or inconsistency). There are some mathematical formulations available for computing intrajudge inconsistency index (e.g., Chang, 1999 for an Angoff method; Chang, Linden, & Vos, 2004 for Interdependent Evaluation of Alternatives method). On the other hand, for interjudge consistency, intraclass correlation coefficient is often used as a measure of interjudge reliability (Kane, 1994). However, no procedure has been developed for computing the interjudge inconsistency index for Body of Work (Kinston, Kahl, Sweeney, & Bay, 2001), Bookmark (Mitzel, Lewis, Patz, & Green, 1986), and a Yes/No variation of Angoff methods (Plake, Ferdous, Impara, & Buckendahl, 2005).

The purpose of this study is to derive a generalized mathematical formulation for interjudge inconsistency index that can be used for the Body of Work, Bookmark, and Yes/No variation of Angoff method. By overall interjudge inconsistency, we are referring to percent of disagreement out of all possible disagreements among the judges ratings either for the test items or for given number of students' body of work.

Method

Data

In this study, three standard setting data sets are used. The first standard setting study was administered for setting cutscores for grade 11 Mathematics in a Midwestern state using a Yes/No variation of Angoff method (Impara & Plake, 1977); second standard setting study was conducted for grade 7 Reading in a Northeastern state using a modification of Bookmark procedure (Lewis, Mitzel, & Green, 2001); and third standard setting was also implemented in a Northeastern state for grade 8 writing using a Body of Work (Kingston, Kahl, Sweeney, & Bay, 2001) method. In each standard setting study, students were to be classified into four achievement level categories (ALCs). Although the achievement level categories were different for different states we are unifying them into four common categories: ALC1, ALC2, ALC3, and ALC4. ALC1 is the lowest and ALC4 is the highest achievement level categories.

Each of these methods was implemented using two rounds. Overall student performance data (item difficulty values, or proportion of students answering each item correctly), impact data (i.e., cumulative percent distribution) and, for the Body of Work method, student profile data (i.e., response pattern of a typical student at different score points comprising the whole raw-score range) were supplied to the judges between the rounds. Before judges started working on the operational tasks, they were given training and adequate time to discuss the performance level descriptors in the group.

In a Yes/No variation of Angoff method, judges are asked to designate an ALC to each item on the test that they think a just ALC student would have a 0.67 or higher probability to answer the test item correctly. In the Bookmark procedure judges put a bookmark just before to the item that they think a just ALC student will have less than 0.67 probability of correct response. In this procedure, items are ordered by item-response theory based difficulty values $b_{0.67}$ after removing guessing from the model. In the Body of Work method, judges classify the selected students' papers into ALCs. In this procedure, students' papers are ordered from low scoring to high scoring and they are often selected using systematic procedures such that the full raw score continuum is covered.

Procedure

The procedure of computing interjudge inconsistency is based on the absolute difference between two judges' responses for all possible of pair of judges. This index can be calculated both at the item-level (or student body of work level) and for the entire test (or set of students body of work). The interjudge inconsistency for an item k is defined as the proportion of total observed inconsistency to the total possible inconsistency. Total observed inconsistency is defined as the sum of the absolute differences of all possible pair of judges' responses.

Interjudge inconsistency for item k is,

$$I_k = \frac{TOI_k}{TI} \qquad (1)$$

$$TOI_k = \sum_{\substack{i=1 \\ l \neq i}}^{\frac{n!}{2!(n-2)!}} | J_{ik} - J_{lk} | \qquad (2)$$

Where,

I_k = Interjudge inconsistency for item k

TOI_k = Total observed interjudge inconsistency for item k

TI = Total possible interjudge inconsistency

n = Number of judges in the standard setting study

J_{ik} = Judge i's response to item k; J = 1, 2, 3, 4 (1= ALC1, ALC4=4)

Total possible inconsistency for each item is,

$$TI = \left[\frac{n!}{2!(n-2)!} - \frac{(n-1)!}{2!((n-1)-2)!} + \frac{(n-2)!}{2!((n-2)-2)!} - \ldots\ldots\ldots - \frac{(n-(n-3))!}{2!((n-(n-3))-2)!} + \frac{(n-(n-2))!}{2!((n-(n-2))-2)!} \right] * d \qquad (3)$$

Where,

d = Maximum absolute possible difference between two judges' ratings. As there are four achievement level categories; one judges may give a rating of 1 (ALC1) to the item and the other judge may give a rating of 4 (ALC4) so possible maximum absolute difference is 3.

Overall interjudge inconsistency for the entire test can be computed by aggregating the item interjudge inconsistency values across the items on the test or across the student paper classification decisions. The overall interjudge inconsistency index can also evaluated by examining its relationship with standard deviation of judges' item performance designations for each item on the test. The overall interjudge inconsistency index for K items on the test is given by,

$$I = K^{-1} \sum_{k=1}^{K} I_k \tag{4}$$

This general procedure (presented above) for computing interjudge inconsistency can be used for a Yes/No variation of Angoff methods without any modifications. However for the Bookmark and Body of Work procedures, this index is calculated with small variations. In Bookmark (items are ordered from easiest to the most difficult) and Body of Work (selected students' papers are ordered based on their raw total scores) procedures, judges "point" on the score scale where they believe the cutpoints should fall. Typically there are large ranges of the score scale where the judges are in agreement. For example, for the Bookmark disagreements most often occur in the vicinity of the cutpoints. Thus, calculating the average interjudge inconsistency across all the items may over estimate their agreement and therefore underestimate their inconsistency. Therefore, interjudge inconsistency in Bookmark and Body of Work procedures is computed only for subset of items in which at least one judge has disagreed with rest of the judges in the panel. We call this "interjudge inconsistency around the cut scores." Thus, interjudge inconsistency for Bookmark and Body of Work procedures is given by,

$$I = \begin{cases} p^{-1} \sum_{k=1}^{p \in K} I_k & I_k > 0 \\ \text{Under Estimated} & \textit{Otherwise} \end{cases} \tag{5}$$

Values of the interjudge inconsistency index lie between 0 and 1, 0 being perfectly consistent and 1 being perfectly inconsistent. Overall maximum possible distance or perfectly inconsistency across the judges may be obtained when half of the judges give one extreme rating (e.g. ALC4) and the other half gives the other extreme value (e.g. ALC1). This is one possible combination of judges' ratings for which the value of the index is 1. This combination of judges' ratings may have very low probability of occurring but it is not impossible. The value of the index depends on two factors: number of judges and number of achievement level categories. As the number of judges increases computation of the index becomes complex.

Illustration

In fictitious yes/no variation of Angoff-based (Figure 1) and Bookmark/Body of Work-based (Figure 2) standard setting studies, a total of 6 judges participated for setting standards ALC1, ALC2, ALC3, and ALC4 for a test of only 5 items. In yes/no method, judges are asked to designate ALCs to each item on the test and for Bookmark/Body of Work method, judges were asked to classify items or students' folders into ALCs. Three cases are presented in the following to demonstrate three situations (Figure 1):

Figures 1 & 2 About Here

In figures 1 and 2, for each case row marginals and the column marginals represent overall inconsistency among the judges (for all pairs of judges together) for item or student's folder k and proportion of total inconsistency due to the pair of judges for the test, respectively. An average across the items (row marginal) and a sum across pairs of judges (column marginal) constitute the overall interjudge inconsistency.

In case 1 (i.e., perfect agreement) when judges gave exactly the same ratings to the items or students' folders, the values of the row marginals, column marginals, and overall interjudge inconsistency value are estimated to 0 (as they were expected).

In case 2 (i.e., a typical situation when judges do not have perfect agreement) the value of the index was estimated to be 0.51 for yes/no method; maximum column marginal and row marginal being 0.06 for the pair of judges 2 and 6 and 0.59 for item 4 respectively. It was also evident from the ratings; judges 2 and 6 did not agree on single item. On the other hand, for Bookmark/Body of work method the value of the index was estimated to be 0.36; maximum column and row marginals being 0.05 (for the pairs of judges 2 and 4, and 4 and 5) and 0.48 for the items or student's folders 3 and 4, respectively. Again, judges 2 and 4 and 4 and 5 did not agree on any single item.

In case 3 (i.e., perfect disagreement or maximum distance) when one half of the judges gave one extreme value ALC1 and the other half gave the other extreme value ALC4 or vice versa. The value of the index was estimated to be 1 (as it was expected).

Results

Results of the study are presented by standard setting methods used. Overall results are appeared to be very consistent with each other. To facilitate comparability, the interjudge indices were calculated both by using all of the items or student papers and then using only the items/papers around the cutscores. These results are shown in Table 1.

Yes/No variation of Angoff: This method was implemented for setting standards for grade 11 Mathematics in a Midwestern state. The test had 105 operational multiple choice items. A group of 14 judges participated in the standard setting study. The cutscore were: 7.0 (SEM=1.1) for ALC1, 42.0 (SEM=5.0) for ALC2 and 84.0 (SEM=3.6) for ALC3. The overall interjudge inconsistency (based on equation 4 using all of the items) was estimated as 0.25. For comparability purposes, this index was also calculated for the items around the cutscores. In order to do this, the items first needed to be ordered by difficulty (using item p-values). Then the individual panelists' cutscores were located on this continuum. Then, the interjudge consistency value around the cutscores was computed. This value did not differ very much from the value obtained from using all of the items; the index using the items around the cutscores was 0.28 as compared to 0.25 when all the items were used in the calculation.

Table 1 About Here

Bookmark method: The Bookmark method was used for grade 7 Reading standard setting study in a Northeastern state. The test had 40 (both multiple choice and constructive response) items. A group of 12 judges participated in the standard setting study. The cut scores were computed as 22.0 (SEM=2.99) for ALC1, 29.0 (SEM=2.86) for ALC2, and 43.0 (SEM=2.35) for ALC3. For only 5 items (out of around 40), at least one judge disagreed with rest of the panel; 3 were for ALC1, 2 were for ALC2, and none for ALC3. The overall interjudge inconsistency around the cuts (using equation 5) was estimated to be 0.14. This index was also calculated using all of the items for comparison purposes. The value of the interjudge inconsistency when all of the items were used equaled to 0.06, substantially smaller than the value derived when using the items around the cuts.

Body of Work: This method was implemented for setting standards for grade 8 Writing in a Northeastern state. A total of around 35 students' papers were used. A panel of 14 judges participated in the standard setting study. The cut scores were: 18 (SEM=1.80) for ALC1, 24 (SEM=1.84) for ALC2, and 32 (SEM=2.41) for ALC3. The overall interjudge inconsistency index using equation 5 was estimated as 0.13. Again, for comparison, the interjudge inconsistency index was computed using all of the data points, not just the ones around the cuts.

The values using all of the data points equaled 0.07. As was seen with the Bookmark method, this value was smaller than the one derived from using only the data points around the cuts.

Discussion

This study examined two approaches to calculating an index of interjudge inconsistency. In one approach, all of the products (items or student papers) that are evaluated were used. For the Angoff and Bookmark methods, this entailed using all of the test questions. For the Body of Work method, this involved using all of the student papers that were used in the judgmental process. All of the panelists' ratings were therefore used in this case when computing the index of interjudge inconsistency.

In the second approach, only the items (or student papers) that were in the vicinity of the cutpoints were used to compute the interjudge inconsistency index (called the interjudge inconsistency around the cuts). The second method is recommended specifically for methods such as the Bookmark and Body of Work where panelists are presented with information about the difficulty of the items (for the Bookmark method) or the performance of the students papers (Body of Work). In those instances, panelists will appear more consistent due to the extra information they are presented to complete their ratings.

For the Bookmark method, there were blocks of items between the location of panelists' bookmarks that will be common across panelists (especially at the beginning and end of the ordered items booklets). Similarly, when panelists are presented student papers in performance order, groupings of student papers will also be in common across panelists' classification decisions. Again, this may be most pronounced at the lower and higher performance levels for student papers. When these indices were computed using both approaches, as expected the two values differed greatest for the Bookmark and Body of Work methods. The agreement level differed the most for the Bookmark method where the index when using all of the items showed nearly perfect agreement (0.06) but had a higher value (indicating higher levels of disagreement) when the index was calculated using the items around the cuts (0.14).

Standard setting methods are being used to make critically important decisions for examinees. These examinees may be given licenses to practice a profession, qualify for high school graduation, and given special awards or privileges based on the outcome of a standard setting procedure. Therefore, the quality of the standard setting study should be evaluated when applying the information derived from a standard setting study.

Evidence to support the accuracy and validity of the results of standard setting studies are critically important for a standard setting study, especially for high stakes tests. Among other indicators of quality, an index of panelists' agreement (or disagreement) can be useful in communicating the technical quality of the results. This mathematical formulation of computing interjudge inconsistency (for multiple standard setting methods) not only would help in validating the standard setting procedure but would also be useful for the standard setting facilitators identifying judges (based on Round 1) whose rating are inconsistent with the rest of the panel. Facilitators could provide these panelists with necessary supports (e.g. feedback on understanding performance level descriptor). This procedure also could identify items for which judges were substantially inconsistent and then if necessary, the facilitator could bring those items for discussions among the judges. So, this interjudge inconsistency index could be helpful for completely documenting standard setting procedure and could also help for designing effective standard setting studies.

This approach has the advantage of being generalizable across several standard setting methods. Other methods for computing intra and interjudge consistency (e.g., Chang, 1999;

Chang, et al, 2004) are method dependent, limiting the comparability of the indices across methods. Further, this approach is applicable to methods that rely of panelists' classifications of items/papers into performance level categories, making it useful with several standard setting methods currently used across states for making performance level decision on state-level assessments for NCLB (2002) reporting. Use of methods such as these could strengthen the validity evidence about the results from a standard setting study.

References

Angoff, W. H. (1971). Scales, norms, and equivalent scores. In R.L. Thorndike (Ed.), *Educational Measurement (*2nd ed.) (pp. 508-60). Washington, DC: American Council on Education.

Berk, R. A. (1986). A consumer's guide to setting performance standards on criterion-referenced tests. *Review of Educational Research, 56*, 137-172.

Chang, L. (1999). Judgmental item analysis of the Nedelsky and Angoff standard setting methods. *Applied Measurement in Education, 12*, 151-165.

Chang, L., van der Linden, W. J., & Vos, H. J. (2004). Setting standards and detecting intrajudge inconsistency using Independent Evaluation of Response Alternatives. *Educational and Psychological Measurement, 64(5)*, 781-801.

Cizek, G.J. (2001). *Setting performance standards: Concepts, methods, and perspectives.* Mahwah, NJ: Lawrence Erlbaum Associates.

Council of Chief State School Officers (2001). *State Student Assessment Programs Annual Survey*. Data Volume II. Washington D.C. Author.

Hambleton, R. K., & Powell, S. (1983). A framework for viewing the process of standard setting. *Evaluation & the Health Professions, 6(1)*, 3-24.

Iowa City, IA: ACT, Inc.

Kane, M. T. (1994). Validating the performance standards associated with passing scores. *Review of Educational Research, 64(3)*, 425-461.

Kingston, N. M., Kahl, S. R., Sweeney, K. P., & Bay, L. (2001). Setting Performance Standards Using the Body of Work Method. In G.J. Cizek (Ed.), *Setting Performance Standards:*

Concepts, Methods, and Perspectives, (pp. 219-248). Lawrence Erlbaum Associates, Mahwah, NJ.

Lewis, D. M., Mitzel, H. C., & Green, D. R. (1996, June). *Standard Setting: A Bookmark Approach*. Paper Presented at the Council of Chief State School Officers National Conference on Large-Scale Assessment, Phoenix, AZ.

Livingston, S. A., & Zieky, M. J. (1982). *Passing scores: A manual for setting standards of performance on educational and occupational tests*. Princeton, NJ: Educational Testing Service.

Lunz, M. (1995). Methods *of setting criterion standards for performance examinations*. Unpublished manuscript.

Plake, B. S., Ferdous, A. A., Impara, J. C., & Buckendahl, C. W. (2005). *Setting Multiple Performance Standards Using the Yes/No Method: An Alternative Item Mapping Method*. Paper Presented to the Meeting of the National Council on Measurement in Education, Montreal. Canada.

Reckase, M. D. (2000, June). The evolution of the NAEP achievement levels setting process: A summary of the research and development efforts conducted by ACT.

Sireci, S. G., & Biskin, B. J. (1992). Measurement practices in national licensing examination programs: A Survey. *CLEAR Exam Review, 3*, 21-25.

U.S. Department of Education. (2002). *No Child Left Behind Act of 2001*, Pub. L. No. 107-110, 115 Stat. 1425 (2002).

Table 1

Results of three standard setting studies

Study	Method	No. of judges	Cut Score			Interjudge inconsistency	
			ALC1	ALC2	ALC3	All items	Around the cut
1	Yes/No	14	7.0 SEM=1.10	42.0 SEM=5.00	84.0 SEM=3.60	0.25	0.28
2	Bookmark	12	11.0 SEM=2.99	25.0 SEM=2.86	43.0 SEM=2.35	0.06	0.14
3	Body of Work	14	18.0 SEM=1.80	24.0 SEM=1.84	32.0 SEM=2.41	0.07	0.13

Figure 1

An Illustration of Computing Interjudge Inconsistency Index for Yes/No Method

Case 1: Interjudge Inconsistency = 0

Computation of Interjudge Inconsistency

	J1	J2	J3	J4	J5	J6	12	13	14	15	16	23	24	25	26	34	35	36	45	46	56		I_k
1	4	4	4	4	4	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0
2	2	2	2	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0
3	4	4	4	4	4	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0
4	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0
5	3	3	3	3	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0
							0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	I=	**0**

Case 2: Interjudge Inconsistency > 0 and <1

Computation of Interjudge Inconsistency

	J1	J2	J3	J4	J5	J6	12	13	14	15	16	23	24	25	26	34	35	36	45	46	56		
1	4	3	3	3	4	2	1	1	1	0	2	0	0	1	1	0	1	1	1	1	2		0.48
2	2	4	3	2	3	2	2	1	0	1	0	1	2	1	2	1	0	1	1	0	1		0.52
3	3	2	2	4	3	3	1	1	1	0	0	0	2	1	1	2	1	1	1	1	0		0.48
4	2	4	4	3	3	2	2	2	1	1	0	0	1	1	2	1	1	2	0	1	1		0.59
5	1	1	2	2	2	3	0	1	1	1	2	1	1	1	2	0	0	1	0	1	1		0.48
							0.04	0.04	0.03	0.02	0.03	0.01	0.04	0.04	0.06	0.03	0.02	0.04	0.02	0.03	0.04	I=	**0.51**

Case 3: Interjudge Inconsistency = 1

Computation of Interjudge Inconsistency

	J1	J2	J3	J4	J5	J6	12	13	14	15	16	23	24	25	26	34	35	36	45	46	56		
1	4	4	4	1	1	1	0	0	3	3	3	0	3	3	3	3	3	3	0	0	0		1
2	1	1	1	4	4	4	0	0	3	3	3	0	3	3	3	3	3	3	0	0	0		1
3	1	1	1	4	4	4	0	0	3	3	3	0	3	3	3	3	3	3	0	0	0		1
4	4	4	4	1	1	1	0	0	3	3	3	0	3	3	3	3	3	3	0	0	0		1
5	1	1	1	4	4	4	0	0	3	3	3	0	3	3	3	3	3	3	0	0	0		1
							0	0	0.11	0.11	0.11	0	0.11	0.11	0.11	0.11	0.11	0.11	0	0	0	I=	**1**

Figure 2

An Illustration of Computing Interjudge Inconsistency Index for Bookmark/Body of Work Method

Case 1: Interjudge Inconsistency = 0 — Computation of Inter-Judge Inconsistency

	J1	J2	J3	J4	J5	J6	12	13	14	15	16	23	24	25	26	34	35	36	45	46	56		I_k
1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0
2	2	2	2	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0
3	3	3	3	3	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0
4	3	3	3	3	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0
5	4	4	4	4	4	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0
							0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	I=	**0**

Case 2: Interjudge Inconsistency > 0 and <1 — Computation of Inter-Judge Inconsistency

	J1	J2	J3	J4	J5	J6	12	13	14	15	16	23	24	25	26	34	35	36	45	46	56		
1	1	1	1	2	1	1	0	0	1	0	0	0	1	0	0	1	0	0	1	1	0		0.19
2	1	1	2	2	1	2	0	1	1	0	1	1	1	0	1	0	1	0	1	0	1		0.33
3	2	1	2	3	1	2	1	0	1	1	0	1	2	0	1	1	1	0	2	1	1		0.48
4	3	2	3	4	2	3	1	0	1	1	0	1	2	0	1	1	1	0	2	1	1		0.48
5	4	3	4	4	3	4	1	0	0	1	0	1	1	0	1	0	1	0	1	0	1		0.30
							0.02	0.01	0.03	0.02	0.01	0.03	0.05	0	0.03	0.02	0.03	0	0.05	0.02	0.03	I=	**0.36**

Case 3: Interjudge Inconsistency = 1 — Computation of Inter-Judge Inconsistency

	J1	J2	J3	J4	J5	J6	12	13	14	15	16	23	24	25	26	34	35	36	45	46	56		
1	1	1	1	4	4	4	0	0	3	3	3	0	3	3	3	3	3	3	0	0	0		1
2	1	1	1	4	4	4	0	0	3	3	3	0	3	3	3	3	3	3	0	0	0		1
3	1	1	1	4	4	4	0	0	3	3	3	0	3	3	3	3	3	3	0	0	0		1
4	1	1	1	4	4	4	0	0	3	3	3	0	3	3	3	3	3	3	0	0	0		1
5	1	1	1	4	4	4	0	0	3	3	3	0	3	3	3	3	3	3	0	0	0		1
							0	0	0.11	0.11	0.11	0	0.11	0.11	0.11	0.11	0.11	0.11	0	0	0	I=	**1**