

The Valency Assessment Test (VAT) : Developing and testing a new measurement tool for human bond *

Med Hafsi

Nara University, Department of Sociology

Abstract

The purpose of the present article is to describe the results of two attempts to test the reliability of a revised version of the Reaction to Group situation Test of Nara University (RGST-Nu) which is supposed to assess "valency", or the individual predisposition to combine with and relate to others, a concept developed by Bion and revised by the author. After almost a decade of using RGST-Nu, and owing to the difficulties encountered when scoring it, the author with his colleagues and students was led to the conclusion that the test was in need for further improvement. One of the attempts to improve it was to rewrite the scoring manual so that it can deal with daily Japanese language and its numerous ambiguous expressions. The second attempt was to change the direction of the 8-point scale, from the lowest (point-1) to the highest (point-8) score, and redefine more precisely the content covered by each point of the scale, providing further examples to facilitate the scoring. The third improvement made was a further reduction of the number of stimulus-situation (SS). Using the data and findings from previous studies, and the mean scores, the author selected 25 SS, namely those with the highest mean score. This newly revised test which was referred to as the Valency Assessment Test (VAT) is, unlike RGST-Nu constituted by 5 dependency SS, 5 fight SS, 5 flight SS, 5 pairing SS, and 5 cooperation SS. It is noteworthy that the term "cooperation" is used here instead of "work" used in Stock and Thelen's RGST and the author's RGST-Nu, because it reflects more Bion's thinking. The present article describe the results of two studies designed and conducted to test the reliability of VAT. The purpose of the first study was to test the reliability of the scoring manual and the second one the reliability of the test itself. As a result, the findings of both studies demonstrated the reliability of VAT as a tool for valency assessment.

Key words: Valency, Assesment Test, Reliability, Human Bond, Bion

The core concept of the present study is valency. It was barrowed by the psychoanalyst Wilfred Bion (1961) from chemistry to refer to an individual predisposition to unconsciously combine with his/her peers to give, create, and preserve what he called the "basic assumption group". The term "group" here does not refer to the members constituting the group, but to

* Accepted September 3, 2006.

the mental activity the latter is indulged in. In his widely influential book, *Experiences in groups*, Bion distinguished two kinds of mental activities: The *work group* and the *basic assumption group*. In the case of work group, the group members are united around a real task or a problem, are characterized by a high sense of reality, cooperation, achievement needs and growth. On the contrary, when the group is displaying the latter, it is dominated by three different types of fantasy or unconscious assumptions, namely, the basic assumption of dependency (baD), the basic assumption of fight/flight (baF), and the basic assumption of pairing (baP) (Hafsi, 1997). Discussing in details these assumptions goes far beyond the scope of the present study. Therefore, it suffices to say that although these assumptions are fantasies, the group behaves "as if" they are real, rational, and agreed upon by the whole group. According to Bion, these assumptions are the result of the group members massive regression to Klein's (1946) early psychotics (paranoid/schizoid and depressive) positions, and their instantaneous and unconscious combination through individual valencies. Continuing Bion's work, other researchers have added new basic assumptions (see Anzieu, 1984; Turquet, 1985; Lion & Gruenfel, 1993).

Bion (1961) argued that a person "can have...no valency only by ceasing to be, as far as mental function is concerned, human" (p. 116), suggesting also that there are individual differences in terms of nature and intensity of the valency. Moreover, there are as much valency types (dependency, fight/flight, and pairing) as basic assumption groups. As developed later by the author (Hafsi, 2006) a person has only one dominant or active valency (ACV), and a three relatively less dominant or auxiliary valency (AXV) types. For instance, a person can have a dependency valency as his/her ACV, and fight/flight, and pairing valencies as his/her AXV, or less dominant valencies. Whether ACV allows him/her to contribute to the group's basic assumption corresponding to it (Bion, 1961), and, according to the author, establish stable interpersonal relationships, the AXV, helps him/her to adjust to his social environment and its different interpersonal situations and conditions (Hafsi, 2006).

Bion did not provide a description of these different valencies. This task was undertaken rather by those researchers who have applied his ideas (e.g., Stock & Thelen, 1958; Thelen, 1954; Armelius & Armelius, 1982; Karterud & Foss, 1989; Hafsi, 1997; 2006). Before briefly adumbrating these valency types, it is noteworthy that these researchers share the idea that the fight/flight type includes two different valencies, which led them to distinguish four valency types, dependency valency (DV), fight valency (FV), flight valency (FIV), and pairing valency (PV).

Briefly speaking, a person with a dominant DV is characterized by a need to rely on and drive others to rely on him/her, and other consequent characteristics such as, a strong preference for vertical interpersonal relationships, low self-evaluation, high trust of and

confidence in others (Hafsi, 2006). The FV is behaviorally characterized by outspokenness, assertiveness, tendency for competitiveness, and consequently a strong preference for debate as a means of interpersonal communication, and other related features. The most prominent characteristics of a person with a FIV is his/her strong tendency to evade conflictual situations that may jeopardize his/her interpersonal relationships, hypersensibility to human relationships, and the consequent need to keep emotional (and sometimes even physical) distance between him/herself and others to preserve the relationship with them. As to the PV, it is characterized by a strong tendency to invite and appeal to, and at the same time, convey and encourage intimate relationships, and, therefore, by a preference for situations that promote and satisfy this tendency.

Measurement of valency

The concept of valency, as originally defined by Bion (1961), comprises, besides the vagueness of its definition, another obstacle that had to be overcome, namely the problem of how to assess or measure it. The first attempt to resolve this problem was made by Stock and Thelen (1958), and their students at the National Training Laboratories in Bethel, Maine. Based on Bion's early work on groups, they developed a sentence complexion test, they called the "Reaction to Group Situation Test (RGST)". This test is composed of 44 group (stimulus-) situations; with each situation describing one type of emotional interaction (fight, flight, dependency, pairing) or work interaction. The test serves to determine the subject's valency through his/her reaction to each of these situation. The subject's reaction is thus scored resorting to a very complicated scoring procedure. Discussing this procedure in details goes far beyond the scope of the present study, therefore we refer the reader to Stock and Thelen's own work and to the author's work (Hafsi, 1997) for a detailed description of and discussion about this initial procedure.

Most of those who have used RGST as developed by Stock and Thelen (1958) and their colleagues have pointed out the difficulty in using and working with the original manuals and scoring procedure (Karterud & Foss, 1989). Using the original procedure to rate even a single protocol takes a long time, and is not therefore appropriate for statistical research. As remarked by a number of researchers (Hare, 1973; McGrath, 1984; Lion & Gruenfeld, 1993; Hafsi, 1997), that is one of the reasons why RGST was not widely known and used.

In an attempt to develop a Japanese version of RGST, the author (Hafsi, 1997) has first reduced the number of items from 44 to 28, and translated them into Japanese. Then he, with the collaboration of students of his, developed a scoring manual, and simplified the scoring procedure by 1) omitting one dimension, clarity (see Stock and Thelen, 1985) which did not

seem indispensable to him; 2) combining the remaining two dimensions (acceptance and manner) into an 8-point scale (point 1 = acceptance with positive action; point 2 = acceptance with negative or ambiguous action; point 3 = acceptance with emotion; point 4 = acceptance with ideation; point 5 = non-acceptance with ideation; point 6 = non-acceptance with emotion; point 7 = non-acceptance with negative or ambiguous action; point 8 = non-acceptance with positive action). Individual written responses or reactions are thus scored from 1 to 8, depending on whether the content of the stimulus-situation is accepted (acceptance) or rejected (non-acceptance), and on the manner (behavioral, emotional, or ideational) with which the subject reacts to it. This translation and simplification gave birth to the RGST-Nu, where Nu stands for Nara University.

A study designed to test the reliability and validity of the RGST-Nu and its scoring manual was then conducted (Hafsi, 1997). The results demonstrated the validity of the scoring manual and the reliability of this test. Consequently, the author, his colleagues and students, have been, since then, using the test to conduct empirical research on the relationship between valency and a large number of individual as well as group behaviors and attitudes, and (psycho-somatic) pathology (Hafsi, 2004; Katsuhara, 2005). However, albeit the fact that those studies have helped us to 1) understand and develop further the concept of valency, and 2) shed light on the effect of valency on several aspects of individual and group behaviors, and 3) has resulted in a considerable methodological progress, they have led us also to the conclusion that RGST-Nu is still in need for further revision and changes.

Towards a New Measurement Tool

The first attempt to change RGST-Nu was to alter its name. As mentioned previously, like RGST, the latter refers to a test measuring the person's Reaction to Group Situation. For it is based on Bion's early definition of valency as a means to relate and contribute to the group's basic assumption group (Bion, 1961). Since the development of RGST-Nu, the author's several re-readings of Bion's work, and his experience with groups, have led him to the conclusion that 1) Stock and Thelen's (1985) RGST was not only based on a limited understanding of Bion's concept of valency, but does not also reflect what it is supposed to measure. Based on this evaluation of RGST, and a new redefinition and further development of the valency concept (Hafsi, 2006), the author has decided to change RGST-Nu into VAT, or Valency Assessment Test.

Table1. VAT stimulus-situation in Japanese and English

Dependency Stimulus-situation

- Q7. リーダーが太郎を助けようとしたとき、太郎は
Q7. When the leader offered to help him, Taro.....
- *Q12. グループがうまく機能しなくなったとき、太郎は
*Q12. When the group was unable to function properly, Taro.....
- Q19. 太郎が「グループには助けが必要だ」と言ったとき、次郎は
Q19. When Taro said the group needs help, Jiro.....
- Q23. リーダーが太郎を助けようとしたとき、次郎は
Q23. When the leader offered to help Taro, Jiro.....
- Q25. グループがうバラバラになりそうだと感じた太郎は
Q25. When the group seemed to be breaking up, Taro.....

Fight Stimulus-situation

- Q4. 太郎に対して腹が立っていることに気付いた次郎は
Q4. When Jiro realized he was angry at Taro, he.....
- Q6. グループが太郎の意見をけなしたとき、太郎は
Q6. When the group disparaged Taro's idea, he.....
- Q11. 多くの人がお互いのあら探しをしていると気付いた太郎は
Q11. When Taro realized people in the group were taking digs at each other, he.....
- Q18. グループに対して腹が立っていることに気付いた太郎は
Q18. When Taro felt hostile to the group, he.....
- Q22. 太郎がグループを非難したとき、次郎は
Q22. When Taro attacked the group, Jiro.....

Flight Stimulus-situation

- Q1. 次郎が冗談を言い出したとき、グループは
Q1. When Jiro was joking, the group.....
- Q8. 幾人かのメンバー討論に参加しなくなったとき、太郎は
Q8. When several members dropped out of the discussion, Taro.....
- Q13. 太郎がボーッとしているように見えたとき、次郎は
Q13. When Taro seemed to be daydreaming, Jiro.....
- Q20. 太郎がミーティングの途中で帰ったとき、グループは
Q20. When Taro left the meeting early, the group.....
- Q24. グループ内の皆が自分勝手な行動をとっていると感じた太郎は
*Q24. When Taro felt people in the group were behaving as they like, he.....

Pairing Stimulus-situation

- Q3. グループ内の特に幾人かのメンバーに親しみを感じたとき、太郎は
Q3. When Taro liked some members more than others, he.....
- Q5. 太郎と花子が20分遅れてきたとき、グループは
Q5. When Taro and Hanako arrived twenty minutes late, the group.....
- Q10. 太郎と次郎は
Q10. Taro and Jiro.....
- Q15. 太郎が私の方に振り向いたとき、私は
Q15. When Taro turned to me, I.....
- Q17. グループ内の1人だけがちやほやされているとき、太郎は
Q17. When the group was particularly friendly towards one of its member, Taro.....

Cooperation Stimulus-situation

- Q2. 太郎が「問題に取りかかろう」と言ったとき、私は
Q2. When Taro said: "Let's get to the problem", I.....
- Q9. 太郎が、グループの能力を測るよう提案したとき、私達は
Q9. When Taro suggested that the group assess its own resources, we.....
- Q14. グループは提案されたやり方を試したかったので、太郎は
Q14. Since the group wanted to test the suggested procedure, Taro.....
- Q16. 次郎が「お互いの気持ちを知る必要がある」と言ったとき、太郎は
Q16. When Jiro said we needed, more information about how we felt, Taro.....
- Q21. 太郎がグループに、問題の根源を考えるように進めたとき、私は
*Q21. When Taro recommended that the group consider the roots of the problem, I.....

The second change consisted in reevaluating the 8-point scale of test. In the case of RGST-Nu, the scale ranged from point-1 as the highest score to point-8 as the lowest one. In order to avoid any misunderstanding, and for the sake of clarity, the scale was thus changed so that point-1 will correspond to the lowest, and point-8 to the highest score.

The third change made was a further a reduction of the number of stimulus-situation (SS). Using the data and findings from previous studies, and the mean scores the author selected 25 stimulus-situations, namely those with the highest mean score. As indicated in Table 1, VAT is thus constituted by 5 dependency SS, 5 fight SS, 5 flight SS, 5 pairing SS, and 5 cooperation SS. It is noteworthy that the term "cooperation" is used here instead of "work" used in Stock and Thelen's (1985) RGST and the author's RGST-Nu, because it reflects more Bion's thinking. For Bion (1961) opposes valency to cooperation, considering the latter as an essential prerequisite for the mental activity he called the work group. Therefore, using the term work instead of cooperation may be misleading.

The fourth major change concerns the scoring manual. After several years of experience with of the RGST-Nu, some colleagues and students have reported a number of difficulties when scoring ambiguous reactions to different SS constituting the test. Since most of these difficulties were due to the ambiguous and ambivalent character of Japanese language itself, it was indispensable to remedy to this problem by 1) reexamining those ambiguous reactions and other possible ones not encountered yet, and by 2) consequently, improving the scoring manual so that the rater is provided with further examples of these ambiguous reactions and the way to rate them. The following two studies, whose methodological aspects and results are described below, were conducted to test the reliability of VAT.

STUDY 1

The purpose of this study was to test the reliability of the newly devised (VAT) scoring manual, using the following method.

1. Method

Participants: A total of 211 (73 women and 138 men) first-year undergraduate students enrolled in an introductory psychology course at Nara University participated in the present study. All participants received a partial course credit for their participation.

Procedure: The participants were collectively administered VAT during the class. A graduate student supervised by the author read each of the 25 stimulus-situation (SS) (see Table 1) presented in a form of a short sentence, allowing an interval of ten seconds in between. The participants were told that there were no "right" and "wrong" answers, and

asked, as mentioned in the face shift of the test, to complete the sentence as soon as possible, in a free association manner, without thinking deeply. For the present test is based on the premiss that prompt and emotional responses reflect more deeply personality traits and aspects than cognitive ones.

To test the validity of the newly devised scoring manual, first a number of 10 VAT protocols were randomly selected from the total 211 protocols. Then, three graduate students, who were trained during two 90-minute sessions, were individually asked to score these protocols using the new scoring manual. The scores attributed to each SS by the three raters were then compared.

Table 2. Comparison of the 3 raters regarding the dependency stimulus-situations using ANOVA

| Protocol Number | Rater's Score | | |
|-----------------|---------------|-----------|-----------|
| | Rater 1 | Rater 2 | Rater 3 |
| P1 | 2.8 (.63) | 2.6 (.67) | 2.6 (.88) |
| P2 | 4.0 (.63) | 4.0 (.63) | 3.8 (.88) |
| P3 | 2.6 (.63) | 2.6 (.67) | 2.6 (.75) |
| P4 | 2.8 (.68) | 2.8 (.67) | 2.8 (.97) |
| P5 | 5.8 (.63) | 5.8 (.88) | 5.8 (.67) |
| P6 | 6.8 (.67) | 6.8 (.79) | 5.8 (.82) |
| P7 | 5.8 (.63) | 5.0 (.67) | 5.0 (.88) |
| P8 | 6.8 (.63) | 6.8 (.63) | 6.8 (.67) |
| P9 | 3.0 (.63) | 3.0 (.67) | 3.3 (.88) |
| P10 | 4.8 (.63) | 5.0 (.67) | 5.0 (.88) |

Note: The values represented here are means and SDs (in parentheses). No statistically significant difference was found.

2. Results

In order to test the inter-rater reliability of the scoring manual, the mean scores attributed by each rater to each of the 25 SS constituting VAT in the 10 randomly selected protocols were compared. First, the mean scores of each rater in each of the valency type (dependency, fight, flight, and pairing) were computed, then the three raters were compared, using a one-way ANOVA. As indicated in Tables 2 through 6, the results revealed no significant inter-rater difference regarding the four SS categories (dependency, fight, pairing, and flight) and cooperation category. The three raters attributed almost equal mean scores to the ten randomly selected protocols in the case of dependency (Table 2), fight (Table 3),

Table 3. Comparison of the 3 raters regarding the fight stimulus-situations using ANOVA

| Protocol Number | Rater's Score | | |
|-----------------|---------------|-----------|-----------|
| | Rater 1 | Rater 2 | Rater 3 |
| P1 | 4.8 (.73) | 4.8 (.73) | 4.8 (.73) |
| P2 | 5.0 (.63) | 5.0 (.63) | 5.0 (.63) |
| P3 | 6.0 (.67) | 6.0 (.67) | 5.8 (.88) |
| P4 | 4.0 (.53) | 3.6 (.67) | 4.0 (.53) |
| P5 | 2.8 (.63) | 2.6 (.67) | 2.6 (.88) |
| P6 | 2.8 (.63) | 2.6 (.67) | 2.6 (.78) |
| P7 | 2.6 (.63) | 2.6 (.63) | 2.4 (.78) |
| P8 | 4.8 (.73) | 5.0 (.67) | 5.0 (.63) |
| P9 | 5.8 (.63) | 5.8 (.67) | 5.6 (.58) |
| P10 | 2.8 (.63) | 2.6 (.67) | 2.6 (.88) |

Note: The values represented here are means and SDs (in parentheses). No statistically significant difference was found.

Table 4. Comparison of the 3 raters regarding the pairing stimulus-situations using ANOVA

| Protocol Number | Rater's Score | | |
|-----------------|---------------|-----------|-----------|
| | Rater 1 | Rater 2 | Rater 3 |
| P1 | 5.8 (.63) | 5.6 (.67) | 5.6 (.73) |
| P2 | 4.8 (.73) | 4.8 (.63) | 4.6 (.73) |
| P3 | 6.8 (.53) | 6.8 (.67) | 6.6 (.67) |
| P4 | 3.8 (.63) | 3.6 (.58) | 3.6 (.58) |
| P5 | 4.0 (.58) | 4.0 (.58) | 4.0 (.58) |
| P6 | 5.2 (.63) | 5.2 (.63) | 5.4 (.58) |
| P7 | 5.0 (.58) | 5.0 (.67) | 5.0 (.73) |
| P8 | 4.0 (.63) | 3.8 (.67) | 4.0 (.58) |
| P9 | 4.2 (.63) | 4.0 (.67) | 4.2 (.63) |
| P10 | 5.6 (.63) | 5.6 (.63) | 5.0 (.58) |

Note: The values represented here are means and SDs (in parentheses). There were no statistically significant differences.

Table 5. Comparison of the 3 raters regarding the flight stimulus-situations using ANOVA

| Protocol Number | Rater's Score | | |
|-----------------|---------------|-----------|-----------|
| | Rater 1 | Rater 2 | Rater 3 |
| P1 | 6.0 (.53) | 6.2 (.67) | 6.4 (.43) |
| P2 | 6.4 (.63) | 6.6 (.67) | 6.6 (.67) |
| P3 | 4.8 (.63) | 4.8 (.63) | 4.8 (.63) |
| P4 | 5.0 (.50) | 5.2 (.60) | 5.0 (.60) |
| P5 | 6.4 (.63) | 6.4 (.63) | 6.2 (.43) |
| P6 | 4.8 (.63) | 4.6 (.60) | 4.6 (.60) |
| P7 | 2.8 (.63) | 2.6 (.67) | 2.6 (.67) |
| P8 | 3.8 (.63) | 3.6 (.67) | 3.6 (.67) |
| P9 | 5.0 (.63) | 5.0 (.63) | 5.0 (.63) |
| P10 | 6.8 (.63) | 6.6 (.60) | 6.6 (.60) |

Note: The values represented here are means and SDs (in parentheses). No statistically significant difference was found.

Table 6. Comparison of the 3 raters regarding the cooperation stimulus-situations using ANOVA

| Protocol Number | Rater's Score | | |
|-----------------|---------------|-----------|-----------|
| | Rater 1 | Rater 2 | Rater 3 |
| P1 | 4.2 (.63) | 4.0 (.60) | 4.0 (.60) |
| P2 | 4.2 (.63) | 4.2 (.67) | 4.0 (.60) |
| P3 | 5.8 (.63) | 5.6 (.67) | 5.8 (.88) |
| P4 | 3.8 (.63) | 3.8 (.63) | 3.8 (.63) |
| P5 | 6.0 (.63) | 6.6 (.67) | 6.6 (.67) |
| P6 | 6.2 (.63) | 6.2 (.63) | 6.0 (.73) |
| P7 | 2.8 (.63) | 2.6 (.67) | 2.6 (.67) |
| P8 | 3.0 (.53) | 3.2 (.56) | 3.2 (.56) |
| P9 | 3.8 (.63) | 3.8 (.63) | 3.8 (.63) |
| P10 | 5.0 (.53) | 5.0 (.53) | 5.2 (.73) |

Note: The values represented here are means and SDs (in parentheses). No statistically significant difference was found.

pairing (Table 4), flight (Table 5) and cooperation (Table 6) stimulus-situations. The findings provided thus a strong support for the reliability of the newly devised scoring manual

STUDY 2

The present study was designed to test the reliability of VAT using test/retest method. The main purpose was thus to examine whether VAT is consistent enough to lead to similar results after it has been administered at two separate times. Unlike other methods used to test the reliability of a test, the test/retest method is, besides being time-consuming, a widely accepted and used method (Patten, 2002; Yu, 2005).

1. Method

Participants: Of the total number (N=211) of students who were administered VAT (study 1), only 53.1 % (N=111; 62 men=55.9%, and 49 women=44.9%) volunteered to participate in the present study which took place one year after the first one. The distribution of the participants in terms of dominant or active valency (ACV) type was as follows: 45 (40.5%) of the participants had a dependency valency, 37 (33.3%) a fight valency, 22 (19.8%) a pairing valency and 7 (6.3%) a flight valency. The participant's dominant or active valency (AC) is determined by, 1) computing the means of the four different valency types, and then 2) selecting the valency type with the highest mean as the participant's AC.

Produce: The participant were gathered and readministered VAT collectively following the procedure described in study 1. The purpose here was to compare the mean scores for each stimulus-situation category (dependency, fight, pairing, flight and cooperation) of the present study with those obtained in study 1, using t-test, expecting no significant change in the participants' scores at each category and consequently in their valency types.

2. Results

As indicated in Table 7, the results of a t-test comparing the means of each stimulus-situation category during the initial test and the retest revealed no statistically significant differences. As expected, the mean scores obtained in the first test were almost identical with those obtained in the second test. With the exception of cooperation, the (Pearson) correlation coefficients between the scores of the test and retest were, as follows, highly significant at the 0.01 level. The correlations coefficients were $r=.648$; $p<.0001$ for dependency; $r=.633$; $p<.0001$ for fight; $r=.775$; $p<.0001$ for pairing; and $r=.930$; $p<.002$ for flight.

Moreover, examining whether there was a change in terms of the participants' valency type, we found that only 3 out of the 111 participants (2.7%) have changed. That is, one

participant changed from dependency to pairing valency, one from fight to flight valency, and another from pairing to dependency. This minor change in valency type is hypothetically understood here as due to the fact that the raters in the test and retest were different persons. Owing to a number of time and space related difficulties it was not possible to have the same persons scoring all the protocols of both the test and retest. Nevertheless, given the non-significant differences found between the two tests, and the small number of people who have changed, we can conclude that the findings of Study 2 provide a strong support for the reliability of VAT.

Table 7. The results of the comparison of the test and retest using t-test

| Stimulus-situation Category | Study 1 | Study 2 |
|------------------------------------|----------------|----------------|
| Dependency | 6.23 (.79) | 6.08 (.82) |
| Fight | 6.10 (.71) | 6.14 (.64) |
| Pairing | 5.43 (.64) | 5.38 (.59) |
| Flight | 5.20 (.59) | 5.05 (.50) |
| Cooperation | 5.36 (1.20) | 5.78 (1.05)* |

Notes: The values represented here are means and SDs (in parentheses).

* Correlation is significant at the 0.01 level (2-tailed).

Conclusion and Discussion

The main purpose of the present article was to describe the finding of two studies designed and conducted to test the reliability of a newly improved version of the author's (Hafsi, 1997) RGST-Nu, renamed here as the Valency Assessment Test (VAT).

In the first study, the author attempted to test the reliability of the scoring manual of VAT. The method consisted in asking three raters to rate separately 10 protocols of the test randomly selected, and then comparing them using ANOVA. The results revealed no significant inter-rater differences, providing thus a strong support for the reliability of the manual. Since the raters had different levels in terms of knowledge an experience of the VAT, the findings suggest that any rater, who has undertaken a minimum training of two 90-

minutes sessions, can accurately rate a VAT protocol.

The second study was conducted to test the reliability of VAT using the test/retest method. That is, the results obtained at two different times were compared using t-test. The study revealed no significant differences between the two periods, and a relatively very minor change in terms of the participants' valency types. These findings show thus that VAT is reliable enough to be used as a measure for valency.

The advantages of VAT, compared with RGST (Stock and Thelen, 1958) and RGST-Nu (Hafsi, 1997) are that it is 1) easier to use; 2) it requires relatively less time for scoring; 3) it does not require of the rater an intensive training; and finally, 4) its manual is more adapted to Japanese language and its numerous ambiguous expressions.

In conclusion, it is noteworthy that although psychoanalytical theories and concepts, like valency for instance, have helped us to deepen our understanding of human bonding, they have not attracted enough attention from non-psychoanalytical researchers. One of the main reason for this is the lack of objective measurement tools (Silverman, 1975; Masling & Schwartz, 1979; Greene & Rosenkrantz, 1986). In this sense, we hope that, as valency is concerned, VAT will contribute to partially remedy this lack that seems to be hindering psychoanalytically-oriented empirical research.

References

- Anzieu, D., 1984. *The group and the unconscious* (B. Kiborne, Trans.). London: Routledge & Kegan Paul.
- Armelius, K., & Armelius, B.A., 1982. *Work and emotionality in small groups working with probabilistic inference tasks* (Report No. 7). Umea, Sweden: University of Umea.
- Bion, W., 1961. *Experiences in groups and other papers*. London: Tavistock.
- Greene, L., & Rosenkrantz, J., 1986. Idiosyncratic needs for fusion and differentiation in groups. In J. Masling (Ed.), *Empirical studies of psychoanalytic theories, Vol. 2* (pp. 197-217). Hillsdale NJ: The Analytic Press.
- Hafsi, M., 1997. Valency and its measurement: Validating a Japanese version of the reaction to group situation test (RGST). *Psychologia*, 40, 152-162.
- Hafsi, M., 2004. A psychoanalytic inquiry into the protomental roots of the sudden increase of tuberculosis in Japan. A bionic interpretation. *Memoirs of Nara University*, 32, 117-133.
- Hafsi, M., 2006. The chemistry of interpersonal attraction: Developing further Bion's concept of "valency". *Memoirs of Nara University*, 34, 87-112.
- Hare, A.P., 1973. *Handbook of small group research*. New York: Free Press.
- Karterud, S., & Foss, T., 1989. The group emotionality rating system: A modification of Thelen's method of assessing emotionality in groups. *Small Group Behavior*, 20, 131-150.
- Katsuhara, C., 2005. Etiology of eating disorder: A Bionic review (in Japanese). *Annual Reports of The Graduate School of Nara University*, 10, 53-76.
- Klein, M., 1946. Notes on some schizoid mechanisms. *International Journal of Psychoanalysis*, 27, 99-110.
- Lion, C.L., & Gruenfel, L.W., 1993. The behavior and personality of work group and basic assumption group members. *Small Group Behavior*, 24, 236-257.
- Masling, J., & Schwartz, M., 1979. A critic of research in psychoanalytic theory. *Genetic Psychology*

Monographs, 100, 257-307

McGrath, J.E., 1984. *Groups: Interaction and performance*. New Jersey: Prentice-Hall, Inc.

Patten, M. L., 2002. *Understanding research methods: An overview of the essentials* (3rd ed.). Los Angeles: Pyczak Publishing.

Silverman, L., 1975. On the role of laboratory experiments in the development of the clinical theory of psychoanalysis. Data on the subliminal activation of aggressive and merging wishes in schizophrenics. *International Review of Psychoanalysis*, 2, 43-64.

Stock, D., & Thelen, H., 1958. *Emotional dynamics and group culture*. Washington, D.C.: National Training Laboratories.

Thelen, H., 1954. *Dynamics of groups at work*. Chicago: University of Chicago Press.

Turquet, P.M., 1985. Leadership: The individual and the group. In A.D. Colman & M. Geler (Eds.), *Group Relations Reader 2*. Washington, D.C.: A.K. Rice Institute.

Yu, C.L., 2005. Test-retest reliability. In K. Kempf-Leonard (Ed.), *Encyclopedia of Social Measurement, Vol. 3* (pp. 777-784). San Diego, CA: Academic Press.