LUND UNIVERSITY

Test-retest reliability of the Shape/Texture Identification testTM in people with chronic stroke

Ekstrand, Elisabeth; Lexell, Jan; Brogårdh, Christina

Link to publication

Test-retest reliability of the Shape/Texture Identification test[TM] in people with chronic stroke

Elisabeth Ekstrand, RPT, MSc[1,2], Jan Lexell, MD, PhD[1,2],

Christina Brogårdh, RPT, PhD[1,2]

From the [1]Department of Health Sciences, Lund University, Sweden, [2]Department of

Neurology and Rehabilitation Medicine, Skåne University Hospital, Sweden

<u>Corresponding author</u>: Elisabeth Ekstrand, Department of Health Sciences,

Physiotherapy, Lund University, Box 157, SE-221 00 Lund, Sweden. E-mail:

elisabeth.ekstrand@med.lu.se. Telephone number: +46 70 486132

Abstract

Objective: To evaluate the test-retest reliability of the Shape/Texture Identification test (STI-test[TM]) in persons with chronic stroke.

Design: A test-retest design.

Setting: University hospital outpatient setting.

Participants: Forty-five persons (mean age 65 years) with mild to moderate impairments in the arm and hand > 6 months post stroke.

Interventions: Not applicable.

Main Measure: The STI-test[TM] was used to assess active touch of the hand. It consists of two subtests: identification of shapes and identification of textures, each in three different sizes. Both hands were assessed twice, one week apart. The reliability of the data was evaluated with weighted Kappa statistics and the Svensson rank-invariant method (percentage agreement, systematic and random disagreements).

Results: The median total score of the STI-test[TM] was 5 points (min-max 0-6 points) for the more affected hand and 6 points (min-max 3-6 points) for the less affected hand at both test occasions. The weighted Kappa coefficient was 0.94 for the more affected hand and 0.55 for the less affected hand. The percentage agreement for the more affected hand was 69% for the subtest shapes and 82% for the subtest textures, and for the less affected hand 62% and 91%, respectively. There were no systematic or random disagreements for any of the subtests.

Conclusion: The STI-test^{TM} is reliable to assess active touch of the hand after stroke.

Key words: active touch; hand; outcome assessment; reproducibility of results;

somatosensory disorders; stroke.

Introduction

About 65% of persons admitted to stroke rehabilitation (subacute and chronic phase)
have remaining somatosensory impairments of the affected arm and hand,[1] which
impact on the ability to use the hand efficiently in daily activities.[2, 3] Somatosensory
function of the hand after stroke is traditionally assessed in clinical settings by the
detection of light touch and proprioceptive position discrimination, which partly are
included in the Fugl-Meyer Assessment of Sensorimotor Recovery After Stroke.[4, 5]
However, tests of light touch and proprioception are poorly standardized and therefore
there are methodological considerations about their accuracy and reliability.[4] Moreover,
assessment of light touch and proprioception are performed in a passive manner
(passive touch), where active hand movements are not permitted. As opposed to being
touched, as in passive assessments of touch, active touch is sensing by touching. In
active touch, tactile and proprioceptive information is integrated during intended hand
movements.[2] Active touch is a prerequisite for exploring and identifying objects'
shapes, textures and sizes. The predominant sensory input to an active hand movement
(for example when grasping and manipulating) is active touch.[6] Active touch has also
been shown to be an essential element in the process of tactile learning after stroke.[7]

One standardized outcome measure that can be used to assess active touch of the hand is
the STI-test[TM] (Shape/Texture Identification test).[8] The test assesses active touch by

4

means of identification of shapes and textures of increasing grades of difficulty. The STI-test[TM] is easy to perform and has shown robust psychometric properties in persons with peripheral nerve injuries and peripheral nerve diseases.[8-11] The test is used in clinical practice to assess somatosensory function after stroke but to the best of our knowledge, no study has evaluated the test-retest reliability in this population. The aim of the present study was therefore to evaluate the test-retest reliability of the STI-test[TM] as an outcome measure of active touch of the hand in persons with chronic stroke.

Methods

This study is a part of a larger data collection evaluating the psychometric properties of outcome measures that assess different aspects of functioning and disability of the arm and hand after stroke.[12, 13] In this study data from the STI-test[TM] are presented.

Persons with stroke were recruited from a university hospital in Sweden during April to December 2013 with the following inclusion criteria: i) at least 6 months post stroke and ii) mild to moderate paresis in their more affected arm and hand (i.e., self-reported weakness, decreased dexterity and/or difficulties to perform daily hand activities but the ability to bring the hand to the forehead and to grasp and release one block of the Box and Block test[14]). Exclusion criteria were: i) inability to understand test instructions due

to impaired cognition and/or communication and ii) other diseases that could affect somatosensory function.

To characterize the participants' arm and hand function, assessments of sensorimotor impairments in both arms and hands were performed. Light touch and proprioception were assessed by the Fugl-Meyer Assessment of Sensorimotor Recovery After Stroke[5], grip strength was measured by the dynamometer Grippit (http://www.catell.se, Hägersten, Sweden)[13], muscle tone was assessed by the Modified Ashworth Scale[15] (classified as present if elbow, wrist or fingers had a score larger or equal to 1), and dexterity was assessed by the Modified Sollerman Hand Function Test[16] (score from 0 to 12 points, where 12 indicates normal dexterity). The less affected arm was assessed before the more affected arm and the sensorimotor assessments lasted about 25 minutes.

Prior to inclusion, information about the purpose of the study was provided and each individual gave his or her written consent to participate. The principles of the Declaration of Helsinki were followed and the study was approved by the Regional Ethical Review Board, Lund, Sweden (Dnr 2012/591).

The Shape/Texture Identification test (STI-test[TM]) (Össur Nordic AB, Uppsala, Sweden, http://www.ossur.se),[8] consist of two subtests: a) identification of three shapes (cube,

cylinder or hexagon) and b) identification of three textures (one, two or three raised metal dots placed in a row) (Figure 1). The test includes three difficulty levels: decreasing sizes of the shapes (15, 8 and 5 mm), and decreasing size and distance between the dots (15, 8 and 4 mm). The score of the STI-test[TM] ranges from 0 to 6 points per hand, and 0 to 3 points (one point for each size) for each subtest. A maximum score of 6 points indicates normal somatosensory function.[8] Excellent test-retest reliability has been demonstrated (weighted Kappa 0.79 to 0.87) in persons with peripheral nerve injuries and diseases, and the minimally detectable change has been defined as > 1.2 points in peripheral nerve injuries.[8, 10, 11]

Active touch of the hand was assessed on two occasions, one week apart, the same day of the week and the same time of the day, in a quiet separate room at the hospital by an experienced physiotherapist (first author) in agreement with the standardized protocol of the STI-test[TM].[8] The participants were seated behind a screen to block their vision, and asked to identify the shapes and textures by active touch (Figure 1). Identification was performed with the pulp of the index finger with instruction not to use the nail. First the shapes of 15 mm were exposed to the less affected hand and then to the more affected hand, followed by the shapes of 8 mm and 5 mm in size. The textures were exposed in the same way. First the largest textures were exposed to the less affected hand and then to the more affected hand. Thereafter, the same procedure was used for the smaller

7

textures. The three shapes and textures for each size were presented randomly to the

participants. To score one point the participants had to identify all three shapes or

textures correctly. When the score was summarized the easier level (larger size) had to

be correct to get one point on a more difficult level (smaller size), otherwise that level

was set to zero. The test took about 10 minutes to complete. The participants were not

informed of their results until they had completed the entire test on each test occasion.

Statistics

Descriptive statistics, frequencies, means and standard deviations (SD) or medians and

minimum and maximum (min-max), were calculated for demographic data and clinical

characteristics of the participants.

The STI-test[TM] scores (ordinal data) from the two test occasions were presented as

frequencies and medians (min-max). The differences of the total sum score (test

occasion two minus test occasion one) were presented as frequencies (number of

persons) and percentage. The test-retest reliability data from the two test occasions were

evaluated both for the total sum score and for the two subtest scores.

The test-retest reliability of the total sum score was evaluated by the Kappa statistics,

the proportion of agreement observed beyond the agreement expected by chance,[17]

using quadratic weights.[18, 19] In the Kappa matrices of the STI-test[TM] seven values (0 to 6) were used for the pairs of the total sum score from the two test occasions. The Kappa coefficients were calculated using the statistical software programme MedCalc, version 15 (http://www.medcalc.org). The strength of the Kappa coefficients was interpreted as < 0.40 poor, 0.40 to 0.75 fair to good, and > 0.75 excellent.[20]

To evaluate the test-retest reliability of the two subtests (shapes and textures), the Svensson rank-invariant method [21] was used. Analyses of the percentage agreement and disagreements (systematic and random) between the two test occasions were performed for each subtest. The systematic disagreement was analysed in relative position and in relative concentration. Relative position explains the degree of systematic change in position (higher/lower) and a positive value indicates that the participants have higher scores on the second test occasion than on the first. The relative concentration expresses the degree of systematic shift in concentration (centred/dispersed) and a positive value indicates that the participants have more centred scores at the second test occasion.[21] Possible values of relative position and relative concentration range from -1 to 1, and zero values indicate a lack of systematic disagreement. The random disagreement (i.e., the variance that cannot be explained by the systematic disagreements) is expressed as the relative rank variance. The relative rank variance ranges from 0 to 1, and the higher the value the more dispersed is the test-retest measurements. The relative position,

relative concentration and relative rank variance values together with 95% confidence intervals (CI) were calculated using a freely available programme developed by Elisabeth Svensson (http://www.oru.se/esi/svensson). Statistically significant values were indicated by a 95% confidence interval that did not cover zero.[21]

Results

Forty-five persons (37 men, 8 women) with chronic stroke and mild to moderate impairments in their arm and hand participated in the study. Their mean age and standard deviation (±SD) was 65 ±7 years and the mean time from stroke onset to first test occasion was 44 ±28 months. The clinical characteristics of the participants are presented in Table 1.

The median total sum score of the STI-test^{TM} was 5 points (min-max 0 to 6) for the more affected hand and 6 points (min-max 3 to 6) for the less affected hand at both test occasions. Thirty-one (69%) participants had somatosensory impairments (total sum score < 6) in the more affected hand, 20 participants (44%) in the less affected hand and 9 participants (20%) in both hands according to the STI-test^{TM}.

Slightly more than 50% of the participants had the same total sum score at test occasion 1 and 2 (i.e., zero point difference) for both hands, but over 90% had at most a 1-point difference in the total sum score (Table 2).

The weighted Kappa coefficient was 0.94 (95% CI 0.90 to 0.98) for the more affected hand and 0.55 (95% CI 0.27 to 0.83) for the less affected hand. The agreement expected by chance was 0.21 for the more affected hand and 0.46 for the less affected hand.

As can be seen in Tables 3 and 4, the scores of the STI-test^{TM} for the more affected hand were more evenly distributed along the scale (from lower to higher values), whereas the scores for the less affected hand were more concentrated to higher values for both subtests. A majority of the participants had only a 1-point difference in scores between the two test occasions and only a few participants had a 2-point difference.

The percentage agreement (Table 5) ranged from 62% to 69% for the subtest shapes and from 82% to 91% for the subtest textures. The systematic disagreements in position and concentration as well as the random disagreements for the more and less affected hand were non-significant (the CI included zero) for both subtests.

Discussion

This study showed that the test-retest agreements of the STI-test[TM] in persons with mild to moderate impairments of the arm and hand after stroke were high for both the more affected hand and the less affected hand, without any significant systematic and random disagreements.

More than 50% of the participants had no difference (zero point) in their total sum score between the two test occasions. As the minimally detectable change in peripheral nerve injuries has been defined as 1.2 points[10] it is reasonable to assume that a difference of at least one point lies within the normal variability also for persons after stroke. Over 90% of the participants had a total sum score of at most 1-point difference, which can be considered as a very high agreement between the two test occasions.

The test-retest reliability, based on the weighted Kappa coefficient for the total sum scores, was excellent[20] for the more affected hand (0.94). This is in agreement with previous reliability studies of the STI-test[TM] in peripheral nerve injuries and diseases.[8, 11] However, the Kappa coefficient for the less affected hand (0.55) was only fair to good,[20] although 96% of the participants had at most 1-point difference between the two test occasions. The explanation for the low Kappa value for the less affected hand could be due to the scores being concentrated to the higher end of the scale (3 to 6 points)

compared to the more affected hand where the whole scale was used (0 to 6 points). The

Kappa coefficient is calculated as the agreement observed beyond the agreement

expected by chance. When few values in the scale are used in a test it is easier to obtain

agreement just by chance.[19] The agreement by chance for the less affected hand (0.46)

was higher than for the more affected hand (0.21) and, consequently, the agreement

observed must be higher to obtain the same Kappa coefficient. This shows that the

Kappa evaluation has limitations when the whole scale is not used, as for the less

affected hand. Furthermore, the Kappa coefficient only evaluates the agreement

between repeated test occasions and not if there are systematic or random

disagreements. However, as previous studies have used the weighted Kappa statistics to

evaluate the test-retest reliability of the total sum score of the STI-test[TM], [8, 11] we also

performed this analysis to enable a comparison of results.

To expand our reliability analysis, we also included analyses of the subtest scores by

means of the Svensson rank-invariant method. The advantage of using the Svensson

method is that it provides information about the agreement between repeated

measurements, but also an understanding of the size and type of disagreement. This

enables the possibility to evaluate if the disagreements are large enough to affect the

test-retest reliability of the measurements. The percent agreement for the identification

of shapes could be considered somewhat low for both hands (more affected hand 69%

and less affected hand 62%) compared to the textures (more affected hand 82% and less affected hand 91%) (cf Table 5).[22] One explanation could be that the identification of shapes require a three-dimensional integration of tactile and proprioceptive information by active touch and is thereby more difficult for the brain to interpret compared to the more two-dimensional identification of textures. However, the dispersion of differences between the tests was small and a majority of the differences were not more than 1 point. Moreover, there were no significant disagreements, neither systematic nor random, for any of the subtests or hands.

When the somatosensory impairments were assessed with the FM-UE test (passive touch), 38% of the participants had impairments in the more affected hand but no one in the less affected hand. When the somatosensory impairments were assessed with the STI-test[TM] (active touch) the corresponding figures were 69% for the more affected hand and 44% for the less affected hand. In clinical practice, light touch and proprioception are usually part of the standard assessment after stroke.[4] However, these tests are performed in a passive manner that is less relevant to grip control compared to active touch.[2] Thus, if only light touch and proprioception are assessed, the somatosensory impairments after stroke might be underestimated.

Several factors can influence the test-retest reliability. In the present study, the test situation was carefully standardized and the test protocol was thoroughly described. A one-week interval between the test occasions was chosen to avoid fatigue, minimize learning effects and to standardize the testing. The participants were tested at the same location, at the same time and day of the week at both test occasions. All participants were in the chronic phase after stroke when spontaneous recovery no longer is expected and therefore considered stable with regard to their somatosensory function.

There is no accepted method to decide the sample size when using the Svensson rank-invariant method.[21] About 30 participants are considered sufficient in test-retest reliability studies of parametrical data,[23] but when non-parametric data is used it is suggested that the sample size should be larger.[24] Therefore, we included 45 participants, which we believe can be considered sufficiently large to evaluate the reliability of the STI-test[TM].

Study Limitations

In the present study, persons with major cognitive impairments or difficulties to communicate were excluded, and more men than women agreed to participate. As the STI-test[TM] assesses active touch, persons with no motor function in their more affected

hand cannot perform the test. Therefore, the results cannot be generalized to the entire stroke population.

As this study was part of a larger data collection comprising assessments of function, activity of the arm and hand, perceived participation and life satisfaction after stroke we had to limit the number of somatosensory outcome measures and variables assessed. We only had three somatosensory measures in the present study and it had been valuable to also include other measures of active touch. However, there is a lack of simple and standardized outcome measures to assess somatosensory function after stroke. In the present study, we therefore aimed to evaluate the psychometric properties of the STI-test™ as a measure of active touch. In the future it would be valuable to investigate other psychometric properties, such as validity and responsiveness, but also how the STI-test™ is associated with motor function (dexterity) of the hand.

Clinical Messages

- The STI-test™ is reliable to assess active touch of the hand in persons with chronic stroke.

Acknowledgements

Funding

Conflict of Interest

No part of this work has been published elsewhere and is not under consideration for publication in any other journal. No conflict of interest exists.

References

1. Carey LM, Matyas TA. Frequency of discriminative sensory loss in the hand after stroke in a rehabilitation setting. J Rehabil Med. 2011;43(3):257-263.

2. Blennerhassett JM, Matyas TA, Carey LM. Impaired discrimination of surface friction contributes to pinch grip deficit after stroke. Neurorehabil Neural Repair. 2007;21(3):263-272.

3. Hermsdorfer J, Hagl E, Nowak DA, Marquardt C. Grip force control during object manipulation in cerebral stroke. Clin Neurophysiol. 2003;114(5):915-929.

4. Winward CE, Halligan PW, Wade DT. Current practice and clinical relevance of somatosensory assessment after stroke. Clin Rehabil. 1999;13(1):48-55.

5. Fugl-Meyer AR, Jaasko L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient. A method for evaluation of physical performance. Scand J Rehabil Med. 1975;7(1):13-31.

6. Gibson JJ. Observations on active touch. Psychol Rev. 1962;69:477-491.

7. Carey L, Macdonell R, Matyas TA. SENSe: Study of the Effectiveness of Neurorehabilitation on Sensation: a randomized controlled trial. Neurorehabil Neural Repair. 2011;25(4):304-313.

8. Rosen B, Lundborg G. A new tactile gnosis instrument in sensibility testing. J Hand Ther. 1998;11(4):251-257.

9. Rosén B, Jerosch-Herold C. Comparing the responsiveness over time of two tactile gnosis tests: two-point discrimination and the STI-test. Brit J Hand Ther. 2000;5(4):114-119.

10. Rosén B. Inter-Tester Reliability of a Tactile Gnosis Test: the STI-Test™. Brit J Hand Ther. 2003;8(3):98-101.

11. Svensson E, Hager-Ross C. Hand function in Charcot Marie Tooth: test retest reliability of some measurements. Clin Rehabil. 2006;20(10):896-908.

12. Ekstrand E, Lindgren I, Lexell J, Brogardh C. Test-retest reliability of the ABILHAND questionnaire in persons with chronic stroke. PM R. 2014;6(4):324-331.

13. Ekstrand E, Lexell J, Brogardh C. Isometric and isokinetic muscle strength in the upper extremity can be reliably measured in persons with chronic stroke. J Rehabil Med. 2015;47:706-713.

14. Mathiowetz V, Volland G, Kashman N, Weber K. Adult norms for the Box and Block Test of manual dexterity. Am J Occup Ther. 1985;39(6):386-391.

15. Bohannon RW, Smith MB. Interrater reliability of a modified Ashworth scale of muscle spasticity. Phys Ther. 1987;67(2):206-207.

16. Rosen B. Recovery of sensory and motor function after nerve repair. A rationale for evaluation. J Hand Ther. 1996;9(4):315-327.

17. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas. 1960;20:37-46.

18. Cohen J. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. Psychol Bull. 1968;70(4):213.

19. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Phys Ther. 2005;85(3):257-268.

20. Fleiss JL. Statistical methods for rates and proportions. 2nd ed. New York: John Wiley; 1981.

21. Svensson E. Different ranking approaches defining association and agreement measures of paired ordinal data. Stat Med. 2012;31(26):3104-3117.

22. Kazdin AE. Artifact, bias, and complexity of assessment: the ABCs of reliability. J Appl Behav Anal. 1977;10(1):141-150.

23. Hobart JC, Cano SJ, Warner TT, Thompson AJ. What sample sizes for reliability and validity studies in neurology? J Neurol. 2012;259(12):2681-2694.

24. Lehmann EL, D'Abrera HJ. Nonparametrics: statistical methods based on ranks. New York: Springer; 2006.

Table 1. Characteristics of the 45 participants with chronic stroke.

Type of stroke, n (%)	
Cerebral infarction	32 (71)
Cerebral haemorrhage	13 (29)
Paretic side, n (%)	
Right	25 (56)
Handedness, n (%)	
Right handedness	42 (93)
Light touch absent or diminished, more affected arm and hand, n (%)[a]	17 (38)
Arm (upper arm, forearm)	12 (27)
Hand (palmar surface)	17 (38)
Proprioception absent or diminished, more affected arm and hand, n (%)[a]	9 (20)
Wrist	5 (11)
Thumb	9 (20)
Grip strength, more affected hand, newton (SD)[b]	238 (112)
Grip strength ratio, more affected hand/less affected hand, newton (SD)	0.71 (0.28)
Spasticity, more affected arm and hand ≥ 1, n (%)[c]	15 (33)
Dexterity (score 0-12), more affected hand, median (min-max)[d]	7 (0-11)

n: number of participants; [a]Fugl-Meyer Assessment of Sensorimotor Recovery After Stroke; [b]assessed by the dynamometer Grippit (www.catell.se, Hägersten, Sweden); [c]Modified Ashworth Scale; [d]Modified Sollerman Hand Function Test.

Table 2. Number (%) of participants with differences in the total STI-test[TM] score between the two test occasions (n=45).

	Differences in scores T2 minus T1				
	-2	**-1**	**0**	**+1**	**+2**
More affected hand	1 (2)	6 (13)	26 (58)	10 (22)	2 (4)
Less affected hand	0 (0)	6 (13)	25 (56)	12 (27)	2 (4)

T1: test occasion one; T2: test occasion two; zero indicates that participants had the same total STI-test[TM] score at both test occasions; a positive or negative value indicates that the ratings at T2 were higher or lower, respectively, than at T1.

Table 3. The subtest shapes: pairs of data from the two test occasions for the more affected hand and the less affected hand (n=45).

		More affected hand				Less affected hand			
		T1				**T1**			
		0	1	2	3	0	1	2	3
	3	-	-	4	**12**	-	1	4	**27**
T2	2	-	2	**6**	3	-	-	-	2
	1	3	**3**	-	1	-	**1**	-	-
	0	**10**	1	-	-	**9**	1	-	-

T1: test occasion one; T2: test occasion two. Numbers in bold print represent the pairs of identical total scores of the STI-test^{TM} at both test occasions.

Table 4. The subtest textures: pairs of data from the two test occasions for the more affected hand and the less affected hand (n=45).

		More affected hand				Less affected hand			
		T1				**T1**			
		0	1	2	3	0	1	2	3
T2	3	-	2	9	**22**	-	-	3	**41**
	2	-	1	**3**	4	-	-	-	1
	1	-	**2**	-	1	-	-	-	-
	0	**1**	-	-	-	-	-	-	-

T1: test occasion one; T2: test occasion two. Numbers in bold print represent the pairs of identical total scores of the STI-test^{TM} at both test occasions.

Table 5. Percentage agreement, and systematic and random disagreements between test occasion 1 and 2 for the two subtests of STI^{TM} (n=45).

	PA %	RP 95% CI	RC 95% CI	RV 95% CI
More affected hand				
Subtest shapes	69	0.03 -0.06 to 0.12	0.05 -0.10 to 0.21	0.02 0.00 to 0.05
Subtest textures	82	0.05 -0.04 to 0.13	-0.09 -0.23 to 0.05	0.00 0.00 to 0.01
Less affected hand				
Subtest shapes	62	0.13 -0.03 to 0.29	-0.03 -0.16 to 0.11	0.07 0.00 to 0.15
Subtest textures	91	0.04 -0.04 to 0.13	0.00 0.00 to 0.00	0.00 0.00 to 0.00

PA: percentage agreement; RP: relative position (systematic disagreement); RC: relative concentration (systematic disagreement); RV: relative rank variance (random disagreement); CI: confidence interval
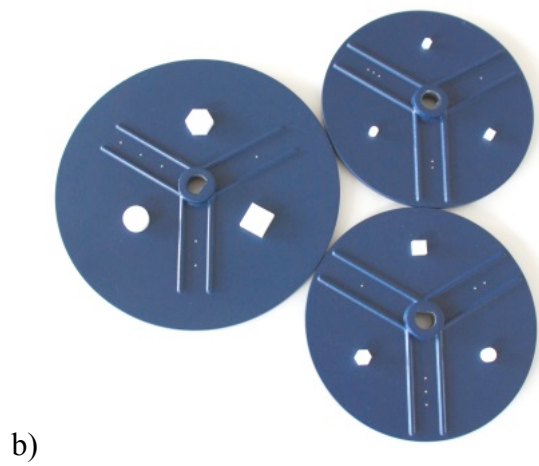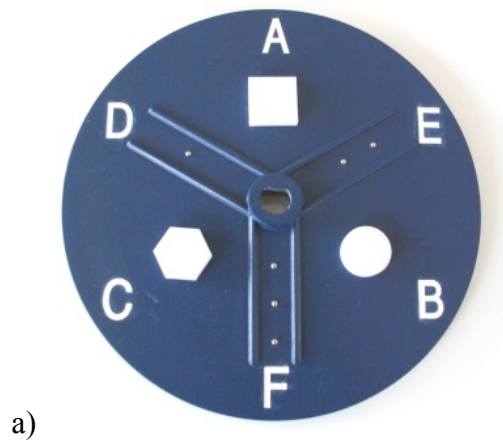
a)

b)

c)

Figure 1: The Shape/Texture Identification test (a-b) and the testing position (c)