

Developing assesment instruments for the sensory acceptability of food products

*¹Sugiyono; ²Djemari Mardapi; ³I Gusti Putu Suryadarma

*Faculty of Engineering, Universitas Negeri Yogyakarta
Jl. Colombo No. 1, Depok, Sleman 55281, Yogyakarta, Indonesia
*Email: giyono@uny.ac.id

Submitted: 05 May 2017 | Revised: 24 May 2017 | Accepted: 24 May 2017

Abstract

The study investigated the constructs and characteristics of physiological, psychological, and sensory instruments. The data were collected by means of a questionnaire, physiological instrument, psychological instrument, and sensory instrument. The expert judgment conclusions were calculated by means of Aiken formula; the instrument construct validity was analyzed with confirmatory factor analysis by the goodness-of-fit test at the significance level of 0.05; the reliability estimate by generalizability with a G-study coefficient ≥ 0.7 and an ICC coefficient ≥ 0.7 ; and the instrument characteristics were analyzed by D-study. The results of the study are as follows: (1) the physiological instrument consisted of nine constructs: four constructs of psychological instrument and five constructs of sensory instrument; (2) three instruments had good face validity, content validity and construct validity, supported by the empirical evidence at $p > 0.05$; (3) the reliability estimate of the three instruments was good and the reliability estimate was supported by empirical evidence with G coefficients of > 0.7 and ICC coefficients of > 0.7 , (4) the three instruments had characteristics that might be appropriate to be used by the university students of culinary program and the vocational high school students of culinary program.

Keywords: *physiological instrument, psychological instrument, sensory instrument*

How to cite item:

Sugiyono, S., Mardapi, D., & Suryadarma, I. (2017). Developing assesment instruments for the sensory acceptability of food products. *REiD (Research and Evaluation in Education)*, 3(1), 92-105. doi:<http://dx.doi.org/10.21831/reid.v3i1.13940>

Introduction

The instruments used in assessing food product from the results of food processing have not been standardized. The assessment is still based on the experience of culinary lecturers and teachers as the instructors. The use of standardized instruments will eliminate rater's subjectivity in assessing the results of the practice. The instruments, as a measurement tool, should have the evidence of validity and reliability (Mardapi, 2008, p. 2).

Food products are used for fulfilling human's basic needs both quantitatively and qualitatively. Food product quantity measurement is conducted by implementing physical measurement tool; on the other hand, the

measurement of food product quality is conducted by implementing sensory acceptability assessment instrument related to human sense responses. The acceptability involves feelings so that the instrument belongs to affective domain.

Food processing is a process of turning the raw material into the product that might be consumed. Food processing for generating food product is conducted by means of certain manners or methods. Food product is used for meeting human's basic needs both quantitatively and qualitatively. The quantity measurement of food product employs the physical measurement tools, while the quality measurement of food product makes use of a

sensory assessment instrument, which is related to human senses.

The measurement of human sense response toward food product involves the principles of sensory sensing, physiological process, and psychological processes (Stone & Sidel, 2004, p. 16). The physiological process should be controlled in order to decrease the physiological error and the psychological process should be controlled in order to decrease the psychological error that might occur. The physiological and psychological error might be controlled by means of physiological and psychological instruments; as a result, the two instruments should be developed.

The physiological aspects consist of attributes of eyes, nose, ears, tongue and skin on the fingers and mouth. The physiological aspects are latent; therefore, the indicators of the attribute measurement should be identified. The attributes of eyes have the measurement indicators of vision and visible light. The attributes of nose have the measurement indicators of odor intensity in the form of olfactory. Then, the attributes of ears have the measurement indicators of ear drum, which is influenced by noise. The sensitivity of tongue is influenced by the temperature of the tested product and the hunger. Last but not least, the skin on the fingers and mouth have the indicators of finger feeling for solid materials and the indicators of mouth feeling for solid or liquid materials.

On the other hand, the psychological aspects that will be involved in sensory measurement are attitude, motivation, adaptation, and situational aspects. The psychological aspects are latent, therefore, the measurement indicators of these attributes should be identified. Attitude is measured through affection or feeling indicators, cognition or knowledge indicators, and conation or tendency indicators. Then, motivation has the measurement indicators of boredom, fatigue and environmental condition. Next, adaptation will influence the sensory benchmark of the measured product and there will be certain changes of response due to the adaptation. Last but not least, situational aspect that might influence the response error of the panelists will be measured by means of errors due to the pre-

testing expectation, errors due to the position of sample presentation, errors due to the stimulus, and errors due to the sample difference, which might be very contrast in the testing sequence.

Psychometric measurement, in addition, should have the stimulus that will be delivered to an individual and then will be processed psychologically in order to generate response. Food product has the attributes that will release the stimulus and these attributes will be sent to an individual. Next, the stimulus will be retrieved by the receptors in human senses and will generate a sensation. The sensation will be sent to the brain and the brain will process the sensation in order to generate perception (Walgito, 1989, p. 54). Based on the perception, there will be responses toward food products under measurement in the form of sensory acceptability. The response measurement toward food products involves feeling that belongs to affective domain.

The measurement scale of affective study that has been widely implemented is Thurstone Scale, Likert Scale and Semantic Differential Scale (Mardapi, 2008, p. 117). Stone and Sidel (2004, p. 8) argue that the interval data scale has been attained from the line-scale measurement scale. Thurstone measurement scale in the form of category data scale is implemented in order to measure interest. Then, Likert scale in the form of ordinal data scale is implemented in order to measure attitude. Semantic Differential Scale is a scale that is limited by the highest and the lowest extreme values. The line scale links (anchors) the lowest and the highest continua. In the line scale, the rater puts a sign on the horizontal line as a reflection of certain sensory characteristic intensity. The measurement scale should be adjusted to the data scale that a researcher demands.

Stone and Sidel (2004, p. 84) state that in order to measure the preference for a product, a researcher might implement hedonic measurement scale technique, for example the nine-point hedonic scale. Watts *et al.* (1989, p. 49) argue that the hedonic scale belongs to the interval data scale. The hedonic scale is very similar to semantic difference scale. The data from the results of sensory testing obser-

vation by means of hedonic scale might be analyzed by implementing parametric statistical test. The hedonic scale should not use the nine-point scale; instead, the hedonic scale might be adjusted to necessity in the sensory testing. A study by Shumate *et al.* (2007, p. 357) has concluded that the number of point scale basically will influence the estimation of generalization. The estimation of generalization increase is in accordance with scale point increase up to a certain level. Up to date, the sensory test instrument refers to Likert Scale for most of the time. However, for conclusion drawing, many researchers refer to the parametric statistical analysis.

The sensory perception might be generated from the process of sensory testing if there are stimuli and receptors. The stimuli come from the attributes of food product and the receptors are possessed by human senses. Perception will occur if the stimulus that has been released by food products is retrieved by the receptors of human senses. The receptors of human senses are unique for they are able to accept certain appropriate stimuli. In addition to having been determined by the stimuli from the attributes of food products, the perception that has been generated is also influenced by the physiological and psychological aspects of the raters.

The objective of the study is to develop a valid, reliable and applicable sensory acceptability assessment instrument of food products, which is expected to be applied as the means for assessing the sensory acceptability of food products resulted from the food processing practice in the culinary study program under the culinary department in vocational high schools and universities.

Method

The study applied research & development (R&D) approach. The stages in the study had been started with field preliminary survey to attain the existing problem of discrepancy as the basis for the stages of development. The developmental stage is conducted by using Borg and Gall (1989) step modification, namely: (1) review of literature and product design planning, (2) preliminary product development, (3) preliminary field

testing, (4) product revision, (5) main field testing, and (6) final product revision.

The research was conducted from December 2014 to March 2015. The study was conducted in two locations, namely: Faculty of Engineering of Universitas Negeri Yogyakarta, and State Vocational High School 4 Yogyakarta.

The population of the study was divided into two: the population from the university and vocational high school. The population from the university was the lecturers and students practicing food processing. On the other hand, the population from the vocational high school consisted of the teachers and students practicing food processing.

The sample of the study was established randomly. Eight lecturers and 36 students of food processing practice course at the Department of Culinary, Faculty of Engineering, Universitas Negeri Yogyakarta and three teachers and 60 students of food processing practice subject at State Vocational High School 4 Yogyakarta were chosen as the sample of the study.

Procedures

Acceptability involves feeling, therefore, it belongs to affective domain. For the development of the affective instrument, the researchers adopted overall steps developed by Mardapi (2008, p. 108): (1) deciding the instrument specification; (2) composing the instrument; (3) deciding the instrument scale; (4) deciding the scoring system; (5) reviewing the instrument; (6) performing experiment; (7) analyzing the instrument; (8) designing the instrument; (9) performing measurement; and (10) interpreting the measurement results. The ten steps were compulsory to perform in developing the affective instrument both for the psychological and physiological attribute measurement and for the learning process. The first step in determining the instrument specification was the substance of the instrument that would be developed. The first step was very important because it was the core of the affective instrument development. The following steps were the elaboration or the follow-up of the previous step.

For the instrument design, the researchers referred to the pattern of instrument specification design developed by Mardapi (2008, p. 109) namely: deciding the measurement objective, deciding the instrument guidelines, selecting the instrument design and format, and deciding the length of the instrument. The instrument generated from the developmental study was to measure the sensory acceptability of food products. The sensory acceptability was measured based on physiological, psychological, and food products aspects. The three attributes were latent and the measurement of the three attributes were based on certain indicators. Then, the researchers designed the operational definition of the three attributes. Based on the operational definition, the instrument guidelines were developed in order to decide the indicator measurement items. The instrument of sensory acceptability was based on the instrument guidelines.

The developmental procedures were performed by modifying the ten steps of R & D approach by Borg and Gall (1989, pp. 784–785) into the following six steps: (1) literature review and product design planning; (2) preliminary product development; (3) preliminary field testing; (4) product revision; (5) main field testing; and (6) final product revision.

The modification performed toward the R & D approach by Borg & Gall is explained as follows. The first and the second step were combined into literature review and product design planning (the first step). Then, the third step was modified into preliminary product development (the second step). Next, the fourth step was modified into preliminary field testing (the third step). Afterward the fifth step was modified into product revision (the fourth step). Further, the sixth step was modified into main field testing (the fifth step). The seventh and eighth steps were omitted by considering the objective of the study. Next, the ninth step was modified into final product (the sixth step). Last but not least, the tenth step was omitted by considering the objective of the study.

The first step, namely literature review and product design planning, included the activities of gathering information and creating the product design plan. Then, the second

step, namely preliminary product development, included the activities of creating a prototype based on the literature review and expert feedbacks. The result of the second step was Prototype-I. Next, the third step, namely preliminary field testing, included the activities of getting evaluation and feedbacks from the experts and practitioners. Afterwards, the fourth and the fifth steps, namely product revision and main field testing, included the activities of evaluating the product and getting the evaluation from the panels of experts. Last but not least, the sixth step, namely field test, included the activities of performing the second step based on the results of the main field testing. The results of the sixth step was a standardized instrument as the final product.

Data, Instrument, Data Gathering and Data Analysis Technique

The data were quantitative with the ordinal data scale. The data were gathered by means of three instruments implemented in the study, namely physiological instrument, psychological instrument, and sensory instrument. In the data gathering activities using the physiological instrument, the respondents as the candidates of sensory acceptability raters should assess themselves regarding their physiological preparedness. Then, the psychological instrument was applied in order to assess their psychological condition by themselves as the rater candidates. If the value that had been attained from the two instruments fulfilled the requirements of the cut value, then the related rater candidates might continue assessing the sensory acceptability of food products by means of the sensory instrument.

Both the test and the non-test instruments should have the evidence of validity and reliability (Mardapi, 2008, p. 15). The face validity was determined through the expert opinion by performing expert judgement, while the content validity was determined by experts through rating activities. The value was decided by calculating the data through the use of Aiken formula:

$$V = \frac{S}{n(c-1)} \quad (1)$$

$$S = \sum s \quad (2)$$

$$s = r - lo \quad (3)$$

Legend:

- V : coefficient of content validity
 S : number of scores from all raters
 n : number of raters
 s : score from each rater
 r : rater's validity rating of items
 c : number of rating category
 lo : lowest validity category
 hi : highest validity category

Note:

- If $lo < hi$ then the formula (3) applies: $s = r - lo$
 If $hi < lo$ then the formula (3) becomes: $s = lo - r$

The content validity of the instrument was calculated by means of Aiken formula and the calculation generated the coefficient of V-count validity. The index of instrument validity was determined by comparing the coefficient of V-count validity and the coefficient of V-table validity. The instrument was regarded meeting the criteria of content validity if the V-count was equal to or bigger than the V-table.

The V-table value was determined by operating the V-Aiken Table on the intersection between the appropriate c column and n line. The 1% level of significance was on the upper line of the same n value and the 5% level of significance was on the lower line of the same n value.

The construct validity was determined by factor analysis with maximum likelihood and referred to the Eigen values ≥ 1 . The data attained from the limited scale experiment were used for attaining the estimation of reliability by means of generalizability theory and Genova analysis with the criterion of coefficient $G \geq 0.7$. Then, the data attained from the expanded scale experiment were used for attaining the estimation of reliability by means of intraclass correlation coefficient (ICC) with the criterion of coefficient $ICC \geq 0.7$. According to (Murthi, 2011, p. 13) and Bartko (1976, p. 763), the formula of ICC correlation coefficient is as follows:

$$r_{ICC} = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_o^2 + \sigma_e^2}$$

Note:

- r_{ICC} : ICC correlation coefficient
 σ_s^2 : subjects' variants (between people)
 σ_o^2 : observer's variant (between people within people)
 σ_e^2 : error variant (residual)

The ICC correlation coefficient might be converted into the Cronbach Alpha correlation coefficient as follows:

$$\alpha = \frac{k \times r_{ICC}}{1 + (k - 1) \times r_{ICC}}$$

Note:

- α : Alpha Cronbach correlation coefficient
 k : number of raters
 r_{ICC} : ICC correlation coefficient

The sensory acceptability assessment instrument was used for assessing food products resulted from the food processing practice. The reference referred by the researchers in the sensory acceptability instrument was criterion reference. The criterion of passing value was determined by means of Standard Setting Method with Extended Angoff (Cizek & Bunch, 2007, p. 88). The cut value calculation was as follows:

$$\text{Cut score} = \text{Means} \times \text{number of items}$$

The self-assessment data using physiological instruments calculated the total score obtained. The result was compared to the cut score of the physiological instruments in the rubric. If the total score \geq cut score, then it can continue the assessment using sensory instruments.

From the self-assessment data gathered by means of the psychological instrument, the total value attained was calculated. The result of the calculation was compared to the cut value of the psychological instrument in the rubric. If the total value \geq the cut value, then the assessment by means of sensory instrument might be continued.

From the data of measurement results by sensory instrument, the mean of the sensory instrument was calculated. The calculation result was compared to the scoring guidelines in the rubric to draw conclusion on the sensory acceptability of food products tested.

Findings and Discussion

The results of the preliminary study show that the implementation of sensory assessment on the results of food processing practice had not made use of a standardized instrument. The assessment was still based on the rater's experience. Then, the literature review shows that the sensory assessment involved physiological and psychological processes (Stone & Sidel, 2004, p. 16).

Physiological process is related to physiological organs that support the activities of sense testing. The sensory organs related to the sensory testing are as follows: sight, olfactory, taste, texture and auditory. The senses retrieve sensory information in the form of stimuli and then the stimuli is converted into electric-loaded nerve impulses that are read by the brain; thus, as a result, sensation occurs (Brennan, 2006, p. 273; Rakhmat, 2007, p. 49). The sensation is retrieved by the receptors in the sensory organs and then sent to the brain. After that, the sensation is returned to the sensory organs as perception. Mason and Nottingham (2002, p. 8) state that stimuli is the chemical or the physical activators that cause responses toward the receptors. For example, human beings have the eye receptor stimulus and sound is the ear receptor stimulus. The receptors are the cells of sensory organs that detect the stimuli, for example the taste buds receptor in the tongue and the light receptor in the retina. The stimuli that has been exposed to the receptors causes sensation which is interpreted psychologically by comparing the sensation to the past experiences; as a result, the sensation is converted into perception. Therefore, the sensory testing might be performed if there is stimuli that has been released by the attributes of food products and the receptors possessed by the senses.

Azwar (2000, p. 5) elaborates that attitude is the constellation of cognitive (mind), affective (feeling) and connative (behavior) components that have interaction from one to another in understanding, feeling and behaving toward an object. An attitude is determined not only by the situation of the object that has been encountered, but also by the past experience, present situation and also

future expectation (Azwar, 2000, p. 3). Attitude itself is an evaluative response. Thus, it is determined by the situational aspect of an individual. Responses would occur only if an individual is encountered by a stimulus that urges a reaction from the individual. Evaluative responses imply that the appearance of the form of reaction that has been stated as an attitude is based on the internal evaluation process of an individual in drawing conclusions from the stimulus in the form of good-bad, positive-negative and also pleasant-unpleasant relationships, then, the individual will potentially react to the object of attitude. The psychological aspects related to the study are the cognitive (mind), affective (acceptance), psychomotoric (adaptation, motivation) and also psychological condition (situational) aspects.

The process of sensory assessment of food products is started from the stimulus that has been exposed to the sense receptors and the stimulus is delivered to the center of nerve constellation in the brain, and as a result, perception occurs (Walgito, 1989, p. 53). The sensing mechanism of food products acceptability assessment includes three stages namely physical process, physiological process and psychological process. The physical process refers to the process of accepting the stimulus that has been released by the object to the sense receptor. The object is in the form of food product and the object releases a stimulus exposed to the sense receptor, and as a result, sensation occurs. The stimulus might come from the external object that has been directly exposed to the sense or from the internal object that has been directly exposed to the sensory accepting nerve. Then, the physiological process refers to the continuation of physical processes. The stimulus that has been accepted by the sense receptor generates sensation and is delivered by the sensory nerve to the center of nerve constellation in the brain, and as a result, response occurs. In other words, the physiological process is the process of delivering the sensation to the center of nerve constellation that generates response. Next, the psychological process refers to the continuation of physiological process. The response that has

been generated in the physiological process is processed in the center of brain nerve constellation, and as a result, perception occurs. Perception is a negotiation between the pattern that people have attained from the environment and the pattern that people have attained from the accumulation of experience (Brennan, 2006, p. 273). Rakhmat (2007, p. 51) states that perception is an experience regarding objects, events and relationships attained by interpreting and also concluding the information from the sensory stimuli. The perception that is attained from the environmental information is combined with the past experience.

In the sensory testing, during the tasting of the food sample, there is an inter-human senses interaction that is resulted in perception. The perception that has been resulted from the stimuli is not purely from one receptor; instead, the perception is the results of an interaction between two or more receptors. The occurrence of the interaction itself has not been surely identified whether it is from the senses as the receptors or from the brain, however, the biggest possibility is that the process of interaction occurs in the brain. Another possibility is that the inter-stimuli interaction comes from the food sample. The perception that has been resulted is not purely from the food, instead, the perception is the result of an interaction between two or more stimuli. The final result of the process is the sensory acceptability of food products.

Based on the literature study and the feedbacks from the experts and practitioners, the instrument guidelines were designed. For the first instrument, the experts who are involved in the study are food science experts, food processing experts and practitioners consisting of food processing lecturers and food processing teachers. The guidelines of the first instrument are clearly presented in Table 1.

Then, for the second instrument, the experts consist of food science experts and psychological experts, while the practitioners consist of culinary lecturers and culinary teachers. The guidelines of the second instrument are clearly presented in Table 2.

Table 1. The physiological instrument guidelines

No.	Attributes	Indicators
1.	Sight	1. Eye health 2. Color blind 3. Lighting condition
2.	Olfactory	1. Nose health 2. Strong odor allergic 3. Smoke allergic 4. Dust allergic 5. Adaptation
3.	Auditory	1. Auditory disorder 2. Auditory sensitivity
4.	Taste	1. Tongue health 2. Taste sensitivity
5.	Texture	1. Teeth health 2. Fingertips skin sensitivity 3. Artificial teeth health

Table 2. The psychological instrument guidelines

No.	Attributes	Indicators
1.	Cognitive	1. Product identification 2. Taste sensitivity 3. Olfactory sensitivity 4. Texture sensitivity
2.	Affective	1. Interestedness in assessing the product 2. Enjoyment of the product 3. willingness to test the product
3.	Psychomotor	1. Adaptation, frequently performing test 2. Adaptation, frequently tasting
4.	Psychological condition	1. Having problems 2. Having physical fatigue (stress)

Table 3. The sensory instrument guidelines

No.	Attributes	Indicators
1.	Appearance	1. Color 2. Presentation 3. Size 4. Form
2.	Smell	1. Easy detection 2. Scent 3. Scent strength 4. Odor, unpleasant scent
3.	Touch	1. Kinesthetic, solid product 2. Consistency, paste product 3. Viscosity, liquid product
4.	Hearing	1. Sweet 2. Salty
5.	Taste	1. Sour 2. Bitter 3. Savory

In the third instrument, the experts consist of food science experts and sensory testing experts while the practitioners consist of culinary lecturers and culinary teachers. The guidelines of the third instrument are presented in Table 3.

The design of sensory acceptability assessment instrument that has been generated from the preliminary study consists of three instruments namely physiological instrument, psychological instrument and sensory instrument. In the preliminary product development stage, the researchers generated the prototype of the instrument, namely the first prototype of the physiological instrument, the first prototype of the psychological instrument, and the first prototype of the sensory instrument. The first prototype instruments were revised in the product revision stage based on the feedbacks and suggestions from the preliminary field test stage. The revision resulted in the second prototype of the physiological instrument, the second prototype of the psychological instrument, and the second prototype of the sensory instrument.

The second prototype of the physiological instrument, the second prototype of the psychological instrument, and the second prototype of the sensory instrument were tested in the main field test stage. These instruments were reviewed based on the feedbacks and suggestions from the state of the main field testing. The result of the revision was the final products. The reliability index of the three instruments was estimated by means of G-Study and of intraclass correlation coefficient (ICC). The validity and reliability of the three instruments were determined before the three instruments were applied as the measurement tools of food product sensory acceptability.

The face validity was determined by means of expert judgement. The experts argued that the instrument had met the face validity.

The content validity was determined by the experts. The values that the experts had provided toward the instruments were calculated by means of Aiken formula. The criterion of content validity was that the V-count coefficient of validity should be equal

or bigger than the V-table coefficient of validity. The V-table coefficient of validity was read on the number of rating category ($c = 4$) and the number of the rater ($n = 8$) and the level of significance of 5% in V-table = 0.75. The results of the calculation showed that the instrument had met the content validity requirement; in other words, the developed instrument was in accordance with the theory. The results of validity calculation showed that 14 out 15 items in the physiological instrument were valid (one item was dropped); 11 items in the psychological instrument were valid; and 17 out of 18 items in the sensory instrument were (one item was dropped).

The physiological instrument was analyzed by means of factor analysis in order to determine the number of factors that should be formed in the instrument. The results of the analysis are presented in Table 4.

Table 4. KMO and Bartlett's test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.653
Bartlett's Test of Sphericity	Approx. Chi-Square	222.321
	df	91
	Sig.	.000

Table 4 shows that the KMO-MSA value is equal to 0.653 and the Chi-Square value is equal to 222.321 with the degree of freedom equal to 91 at the level of significance of 0.000. The KMO value of $0.653 > 0.500$ implies that the developed instrument is in the 'Good' category. The Chi-Square value is equal to 222.321 and the level of significance is Sig. 0.000; in other words, the instrument can be applied for the factor analysis. The results of the factor analysis is that 14 items in the physiological instrument are categorized into five factors (constructs): sight, olfactory, auditory, taste and texture. The factor analysis for determining the construct validity by means of goodness of fit with the maximum likelihood method is presented in Table 5.

Table 5. The goodness-of-fit test

Chi-Square	df	Sig.
19.314	31	.949

The Chi-Square value from the results of calculation is equal to 19.314 on the degree of freedom of 31 at the significance level of 0.949. From the results of analysis, it is found that the p value of 0.949 is bigger than α 0.05; in other words, the constructs developed from the theory are not different from those developed from the analysis on the empirical data. Based on the explanation, it is concluded that the categorization of the items into the factors or the constructs is valid based on the construct validity. Therefore, based on the construct validity with $x^2 = 19.314$ and $p = 0.949$ in the Goodness-of-fit test analysis, the physiological instrument is valid.

The psychological instrument was analyzed by means of the factor analysis in order to determine how many factors should be formed in the instrument. The results of the analysis are presented in Table 6.

Table 6. KMO and Bartlett's test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.			.569
Bartlett's Test of Sphericity	Approx. Chi-Square	149.290	
	df	55	
	Sig.	.000	

Table 6 shows that the KMO-MSA value is equal to 0.569 and the Chi-Square value is equal to 149.290 with the degree of freedom equal to 55 at the significance level of 0.000. The KMO value of $0.569 > 0.500$ implies that the developed instrument is in the 'Good' category. The Chi-Square value of 149.290 at the significance level 0.000 shows that the instrument can be applied for the factor analysis.

The results of the factor analysis show that the 11 items in the psychological instrument are categorized into four factors (constructs): cognitive, affective, psychomotor, and psychological condition factors. The factor analysis to determine the construct validity by means of goodness-of-fit with the maximum likelihood method is shown in Table 7.

Table 7. The goodness-of-fit test

Chi-Square	df	Sig.
18.185	17	.377

The Chi-Square value from the results of the calculation is equal to 18.185 on the degree of freedom of 17 at the significance level of 0.377. The results of the analysis show that p 0.377 is bigger than α 0.050; in other words, the constructs that had been developed from the theory are not different from those developed from the empirical data analysis. Based on the explanation, it can be concluded that the categorization of the items into the factors or the constructs is valid based on the construct validity. Therefore, the psychological instrument is valid based on the construct validity with $x^2 = 18.185$ and $p = 0.377$ in the Goodness-of-fit Test analysis.

The sensory instrument was analyzed by means of the factor analysis to determine the number of the factors that should be formed in the instrument. The results of the analysis are presented in Table 8.

Table 8. KMO and Bartlett's test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.			.652
Bartlett's Test of Sphericity	Approx. Chi-Square	375.062	
	df	136	
	Sig.	.000	

Table 8 shows that the KMO-MSA value is equal to 0.652 and the Chi-Square value is equal to 375.062 with the degree of freedom equal to 136 at the significance level of 0.000. The KMO value of $0.652 > 0.500$ implies that the developed instrument is in the 'Good' category. The Chi-Square value of 375.062 at the significance level of 0.000 showed that the instrument can be applied for the factor analysis.

The results of the factor analysis, namely the 17 items in the sensory instrument, were grouped into five factors (constructs) as follows: sight, olfactory, texture, auditory and taste. The factor analysis for determining the construct validity by means of goodness-of-fit with the maximum likelihood method is presented in Table 9.

Table 9. The goodness-of-fit test

Chi-Square	df	Sig.
56.415	61	.643

The Chi-Square value from the results of calculation is equal to 56.415 with the degree of freedom of 61 at the significance level of 0.643. The results of analysis show that the p value of 0.643 is bigger than the α value of 0.050; in other words, the constructs that had been formed from the theory are not different from those formed from the empirical data analysis. Based on the explanation, it can be concluded that the categorization of the items into the factors or the constructs is valid based on the construct validity. Therefore, the sensory instrument is valid based on the construct validity with $\chi^2 = 56.415$ and $p = 0.643$ in the Goodness-of-fit Test.

For the estimation of instrument reliability in the limited scale experiment, Genova analysis was implemented. The three instruments were tested toward 36 university students majoring in culinary and 60 vocational high school students majoring in culinary. The summary of the results of the reliability estimation by means of Genova analysis is presented in Table 10.

Table 10. Results of G coefficient analysis

Instrument	Facet	Number of Item	G Coeff.	Criteria ≥ 0.70
1. Physiological	Univ. Students	14	0.9288	> 0.7
	VHS Students	14	0.9058	> 0.7
	Mean	14	0.9173	> 0.7
2. Psychological	Univ. Students	11	0.8705	> 0.7
	VHS Students	11	0.8837	> 0.7
	Mean	11	0.8771	> 0.7
3. Sensory	Univ. Students	17	0.9467	> 0.7
	VHS Students	17	0.9602	> 0.7
	Mean	17	0.9535	> 0.7

Table 10 shows that the results of physiological instrument experiment for university students majoring culinary results in the G coefficient value of $0.9288 > 0.7$, which implies that the physiological instrument has met the reliability requirement. The table also presents that the results of physiological instrument experiment for vocational high school students majoring culinary results in the G coefficient value of $0.9058 > 0.7$, which implies

that the physiological instrument has also met the reliability requirement. The mean of coefficient G of $0.9058 > 0.7$ shows that the physiological instrument will meet the reliability requirement if the physiological instrument is applied altogether to the university and vocational high school students majoring in culinary.

Table 10 shows that the results of psychological instrument experiment for the university students majoring culinary resulted in the G coefficient value of $0.8705 > 0.7$, which implies that the psychological instrument has met the reliability requirement. On the other hand, the table shows that the results of psychological instrument experiment for vocational high school students majoring culinary resulted in the G coefficient value of $0.8837 > 0.7$, which implies that the psychological instrument has met the reliability requirement as well. The mean of G coefficient value of $0.8771 > 0.7$ implies that the psychological instrument will meet the reliability requirement if the psychological instrument is applied altogether in the university and vocational high school students majoring culinary.

Table 10 shows that the results of sensory instrument experiment for the university students majoring culinary resulted in the G coefficient value of $0.9467 > 0.7$, which implies that the sensory instrument has met the reliability requirement. On the other hand, the table shows that the results of sensory experiment for vocational high school students majoring in culinary resulted in the G coefficient value of $0.9602 > 0.7$, which implies that the sensory instrument has met the reliability requirement. The mean of G coefficient value of $0.9535 > 0.7$ implies that the sensory instrument will meet the reliability requirement if the sensory instrument is applied altogether in the university and vocational high school students majoring culinary.

For the estimation of instrument reliability in the expanded scale experiment, the researchers implemented intraclass correlation coefficient (ICC) analysis. The three instruments were tested toward 16 university students majoring in culinary and 16 vocational high school students majoring in culinary. The summary of the results of reliability instru-

ment by means of ICC analysis is presented in Table 11.

Table 11 shows that from the testing toward university students majoring culinary, the physiological instrument resulted in the coefficient of reliability of $0.747 > 0.7$, which implies that the physiological instrument is reliable and stable for the university students. On the other hand, the table shows that from the testing toward vocational high school students majoring culinary, the physiological instrument resulted in the reliability coefficient of $0.765 > 0.7$, which implies that the instrument is reliable and stable for the vocational high school students majoring culinary. The mean of reliability coefficient value of $0.756 > 0.7$ implies that the instrument is reliable and stable and it can be applied altogether in the university and vocational high school students majoring in culinary.

Table 11. Results of ICC coefficient analysis

Instrument	Facet	Number of Item	ICC Coeff.	Criteria ≥ 0.70
1. Physiological	Univ. students	14	0.747	> 0.7
	VHS students	14	0.765	> 0.7
	Mean	14	0.756	> 0.7
2. Psychological	Univ. students	11	0.78	> 0.7
	VHS students	11	0.715	> 0.7
	Mean	11	0.712	> 0.7
3. Sensory	Univ. students	17	0.730	> 0.7
	VHS students	17	0.733	> 0.7
	Mean	17	0.732	> 0.7

Table 11 shows that from the testing toward university students majoring culinary, the psychological instrument resulted in the coefficient of reliability of $0.78 > 0.7$, which implies that the instrument is reliable and stable for the university students. On the other hand, the table shows that from the testing toward vocational high school students majoring culinary, the psychological instrument resulted in the reliability coefficient $0.715 > 0.7$, which implies that the instrument is reliable and stable for the vocational high school students majoring culinary.

The mean of the reliability coefficient value of $0.756 > 0.7$ implies that the instrument is reliable and stable and can be applied altogether in the university and vocational high school students majoring in culinary.

Table 11 shows that the sensory instruments tested to the university students majoring culinary resulting in an ICC reliability coefficient of 0.73 exceeding the required minimum 0.7 ICC coefficient criterion, meaning reliable and stable sensory instruments for university students majoring in culinary facets. Similarly, the experiments to vocational high school students majoring culinary produce an ICC reliability coefficient of 0.733 exceeding the criteria of the coefficient of ICC at least 0.7 , which means that it is reliable and stable sensory instruments for students of vocational high school majoring culinary facets. The average coefficient of sensory instrument reliability of 0.732 exceeds the criteria of the coefficient of ICC of at least 0.7 , which means that the sensory instrument is reliable and stable and can be used for either university students or vocational high school students majoring culinary together.

The characteristics of the food product sensory acceptability assessment instrument were related to the number of minimum items that should be involved in the assessment of the criteria of minimum score that should be gained. For deciding the number of minimum item, D-study analysis was implemented, while for deciding the criteria of minimum score, standard setting with the Extended Angoff method was implemented. The summary of the results of analysis is presented in Table 12.

Table 12. Instrument characteristics

Instrument	Facet	Number of items	Min. score	Min. items	Min. score
1. Physiological	Univ. student	14	77.14	≥ 3	16.33
	VHS student	14	77.14	≥ 4	22.04
2. Psychological	Univ. student	11	59.99	≥ 4	20.36
	VHS student	11	59.99	≥ 4	20.36
3. Sensory	Univ. student	17	85.85	≥ 3	15.15
	VHS student	17	85.85	≥ 2	10.10

Table 12 shows that the number of physiological instrument items is 14 items and the criteria of minimum score that should be gained is 77.14. If the physiological instrument is applied toward the university students majoring in culinary, then the number of minimum items will be 3 items and the criteria of minimum score that should be gained would be 16.33. On the other hand, if the physiological instrument is applied toward the vocational high school students majoring in culinary, then the number of minimum items is 4 items and the minimum score that should be gained will be 22.04.

Table 12 shows that the number of psychological instrument items is 11 items and the minimum score that should be gained is 59.99. If the psychological instrument is applied toward the university students majoring in culinary, then the number of minimum items is 4 items and the minimum score that should be gained will be 20.36. On the other hand, if the psychological instrument is applied toward the vocational high school students majoring in culinary, then the number of minimum items will be 4 items and the minimum score that should be gained will be 22.04.

Table 12 shows that the number of sensory instrument items is 17 items and the minimum score that should be gained is 85.85. If the sensory instrument is applied toward the university students majoring in culinary, then the number of minimum items will be 3 items and the minimum score that should be gained will be 15.15. On the other hand, if the sensory instrument is applied toward the vocational high school students majoring in culinary, then the number of minimum items will be 2 items and the minimum score that should be gained will be 10.10.

For deciding the benchmark on the criteria of minimum score, the researchers implemented the standard setting with the Extended Angoff method. The results of the calculation on the criteria of minimum score (cut score) for the physiological instrument is presented in Table 13.

Table 13. Criteria of the physiological instrument score

Number of Item	Cut Score	Minimum Score	Maximum Score
1	5.51	1	7
2	11.02	2	14
3	16.53	3	21
4	22.04	4	28
5	27.55	5	35
6	33.06	6	42
7	38.57	7	49
8	44.08	8	56
9	49.59	9	63
10	55.10	10	70
11	60.61	11	77
12	66.12	12	84
13	71.63	13	91
14	77.14	14	98

The results of the calculation on the criteria of minimum score (cut score) for the psychological instrument is presented in Table 14.

Table 14. Criteria of the psychological instrument score

Number of Item	Cut Score	Minimum Score	Maximum Score
1	5.09	1	7
2	10.18	2	14
3	15.27	3	21
4	20.36	4	28
5	25.45	5	35
6	30.54	6	42
7	35.63	7	49
8	40.72	8	56
9	45.81	9	63
10	50.90	10	70
11	55.99	11	77

The results of the calculation on the criteria of minimum score (cut score) for the sensory instrument is presented in Table 15.

The physiological instrument was applied for measuring the rater's self-assessment toward the physiological preparedness in assessing food product acceptability. For the criteria of minimum score that should be gained in order that the related rater might continue the sensory assessment, sensory instrument was implemented. If the criteria of the minimum

score was not met, then the rater might not continue the process of sensory assessment toward food products resulting from the food processing practice or, under certain conditions, the rater might continue the sensory assessment with the other team of raters.

Table 15. Criteria of the sensory instrument score

Number of Item	Cut Score	Minimum Score	Maximum Score
1	5.05	1	7
2	10.10	2	14
3	15.15	3	21
4	20.20	4	28
5	25.25	5	35
6	30.30	6	42
7	35.35	7	49
8	40.40	8	56
9	45.45	9	63
10	50.50	10	70
11	55.55	11	77
12	60.60	12	84
13	65.65	13	91
14	70.7	14	98
15	75.75	15	105
16	80.80	16	112
17	85.85	17	119

The psychological instrument was applied for measuring the rater's self-assessment toward the psychological condition in assessing food products acceptability. For the criteria of minimum score that should be gained in order that the related rater might continue the sensory assessment, sensory instrument was implemented. If the criteria of minimum score was met, then the rater might not continue the process of sensory assessment toward food products resulting from the food processing practice or, under certain conditions, the rater might continue the sensory assessment with the other team of raters.

The sensory instrument was implemented in order to assess the sensory acceptability of food products resulting from the food processing practice. The results of the assessment were in the form of the scores that ranged from 17 to 119 if all of the 14 items had been implemented. The score that had been attained was interpreted by means of

qualitative score range guidelines from the low extreme to the high extreme. The guidelines regarding the qualitative score were presented separately in the rubric of food product sensory acceptability assessment instrument.

Conclusion and Suggestions

Based on the calculation and analysis of the research data, three conclusions can be drawn as follows: (1) the three instruments meet the face validity, content validity, construct validity requirements; (2) the three instruments meet the reliability requirement based on Genova analysis and ICC analysis; (3) the determination of the assessor to be able to assess the acceptability of sensory food products meets the cut score requirements.

Since the food product sensory acceptability assessment instruments as the result of this research and development have not been disseminated, the instruments should be disseminated in order that they might be applied in the wider scale for meeting the ever-proven validity and reliability.

References

- Azwar, S. (2000). *Sikap manusia: Teori dan pengukurannya* (2nd ed.). Yogyakarta: Pustaka Pelajar.
- Bartko, J. J. (1976). On various intraclass correlation reliability coefficients. *Psychological Bulletin*, 83(5), 762–765. <https://doi.org/10.1037/0033-2909.83.5.762>
- Borg, W. R., & Gall, M. D. (1989). *Educational research: An introduction* (5th ed.). New York, NY: Longman.
- Brennan, R. L. (2006). *Educational measurement* (4th ed.). Westport, CT: Praeger.
- Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests*. Thousand Oaks, CA: Sage Publications.
- Mardapi, D. (2008). *Teknik penyusunan instrumen tes dan nontes*. Yogyakarta: Mitra Cendekia.

- Mason, R. L., & Nottingham, S. M. (2002). *Sensory evaluation manual*. Phitsanulok, Thailand: Naresuan University.
- Murti, B. (2011). Validitas dan reliabilitas pengukuran. In *Matrikulasi Program Studi Doktorat*. Surakarta: Fakultas Kedokteran, Universitas Sebelas Maret.
- Rakhmat, J. (2007). *Psikologi komunikasi*. Bandung: Remaja Rosdakarya.
- Shumate, S. R., Surles, J., Johnson, R. L., & Penny, J. (2007). The effects of the number of scale points and non-normality on the generalizability coefficient: A Monte Carlo study. *Applied Measurement in Education*, 20(4), 357–376. <https://doi.org/10.1080/08957340701429645>
- Stone, H., & Sidel, J. L. (2004). *Sensory evaluation practices* (3rd ed.). Waltham, MA: Academic Press. <https://doi.org/10.1016/B978-012672690-9/50013-5>
- Walgito, B. (1989). *Pengantar psikologi umum* (Rev. ed.). Yogyakarta: Andi Offset.
- Watts, B. M., Ylimaki, G. L., Jeffery, L. E., & Elias, L. G. (1989). *Basic sensory methods for food evaluation*. Ottawa, Ontario: International Development Research Centre.