Routledge
Taylor & Francis Group

International Journal of Science Education

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tsed20

DAST or VoSAL? Adaptation of the VoSAL to Turkish and a comparison between the instruments

Menşure Alkış Küçükaydın & Seher Esen

To cite this article: Menşure Alkış Küçükaydın & Seher Esen (2023): DAST or VoSAL? Adaptation of the VoSAL to Turkish and a comparison between the instruments, International Journal of Science Education, DOI: 10.1080/09500693.2023.2187674

To link to this article: https://doi.org/10.1080/09500693.2023.2187674

Published online: 30 Mar 2023.

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION
https://doi.org/10.1080/09500693.2023.2187674

DAST or VoSAL? Adaptation of the VoSAL to Turkish and a comparison between the instruments

Menşure Alkış Küçükaydın [a] and Seher Esen [b]

[a]Necmettin Erbakan University Eregli Faculty of Education, Konya, Turkey; [b]Selçuk University Faculty of Education, Konya, Turkey

ABSTRACT

The Draw-A Scientist Test (DAST) has been used for many years as a data collection tool in studies where the image of the scientist is attempted to be determined. In this test, in which the image of scientist is attempted to be revealed with drawings, the drawings are analyzed using a coding ruler or checklist. However, the DAST has been criticized in terms of some methodological issues and it has been emphasized that alternative instruments should be developed. In this context, the Views of Scientist, their Activities, and Locations (VoSAL), developed as an alternative to the DAST, are discussed in this study. First, a validity and reliability study was carried out with the adaptation of the VoSAL into Turkish. At this stage, data were collected from 625 pre-service teachers. The analysis results showed that the VoSAL had excellent goodness-of-fit indices. In the second stage of study, the scientist image of 46 pre-service teachers were evaluated comparatively using the DAST and VoSAL. The results showed that the data of the DAST and VoSAL confirmed each other. The results were then discussed in terms of the scientist image literature.

ARTICLE HISTORY
Received 28 October 2022
Accepted 2 March 2023

KEYWORDS
DAST; preservice teachers; scientists; VoSAL

Introduction

Among the main objectives of science education are individuals' having scientific literacy competencies, encouraging career development in science, and gaining a positive scientist image (Finson et al., 1995). However, stereotypical images are encountered in many scientist image studies (Bozzato et al., 2021; Ferguson & Lezotte, 2020; Miller et al., 2018) which began with the studies of Mead and Metraux (1957) and continue today. However, some changes have been encountered in the scientist image studies carried out with pre-service teachers, especially in the last 10 years. According to this, it was reported that there has been a decrease in the images of pre-service teachers about the scientist, which indicates the mythic scientist and technology indicators, and there have also been some changes in the hair and age of the scientist (McCarthy, 2015; Milford & Tippett, 2013; Subramaniam et al., 2013). In addition, it was emphasised that the image of the scientist is related to the cultural context (McCann & Marek, 2016). In this context, it is recommended to investigate the scientist image of pre-service teachers, examine the reasons that lead to image changes, and to adopt new approaches based on

CONTACT Menşure Alkış Küçükaydın mensurealkis@hotmail.com

the findings (McCarthy, 2015). The Draw-A-Scientist Test (DAST) developed by Chambers (1983) has been used for a long time in studies determining the image of the scientist because of its convenience in terms of not containing written instructions and being used with other methodologies (Avraamidou, 2013). However, some studies have reported that the DAST has some methodological issues (Chang et al., 2020; Reinisch et al., 2017). The literature recommends the development of different measurement tools as alternatives to the DAST to overcome these issues (Edgerly et al., 2021; Lamminpää et al., 2020; Reinisch et al., 2017). The Views of Scientists, their Activities, and Locations (VoSAL), which is Likert-type, presented by Reinisch and Krell (2023), as an alternative to the DAST in studies determining the scientist image, is the only known alternative for now. Therefore, in the present study, the findings of the VoSAL on a Turkish sample were presented, and then the data obtained from the DAST and VoSAL were compared.

Theoretical background

Image of the scientist

It is important to determine the image of the scientist in revealing scientific awareness, because the image of scientist in individuals affects the perspective towards science and scientists (DeWitt et al., 2014). Negative attitudes or wrong thoughts towards scientists distance individuals from scientists and science (Christidou et al., 2016). However, there are many factors that affect the image of the scientist. Media, family, peer learning, cultural background, textbooks, and teachers are very important in shaping the image of scientist (Avraamidou, 2013; Bianchini et al., 2000; Farland-Smith, 2009; Tan et al., 2017). Determining the image of the scientist can provide information about the impact of all of these factors. In addition, by detecting the image of the scientist, opinions about the nature of science can be revealed (Sharma & Honan, 2020). Thus, it is possible to have information about the social and cultural dimensions of science, the fact that scientific knowledge is a product of human effort, and the understanding of scientific methods (Christidou et al., 2021). Through this information obtained, strategies can be developed that will expand the image of scientists as individuals, support the development of scientific literacy, and thus contribute to the nature of science (Lamminpää et al., 2020). Science education researchers, curriculum developers and teachers have a great role in the development of these strategies. However, according to Rosenthal (1993), the way educators teach is influenced by stereotypical images of scientists. This situation shows that, it is necessary to investigate the scientist image of pre-service teachers. Because pre-service teachers will transfer their understanding of the nature of science and their attitudes towards science and scientists to their classes (Sharma & Honan, 2020). Accurate and comprehensive measurement tools are needed for research to be conducted in this context.

The DAST and methodological issues

The DAST, developed by Chambers (1983), is widely used as a data collection tool in studies in which the image of the scientist is attempted to be determined. The DAST, which has a very long career, has been used in various parts of the world and the results obtained have confirmed each other many times. Accordingly, it has been observed that individuals mostly have the image of a male scientist with a coat, glasses, working in the laboratory, and messy hair (Alkış Küçükaydın, 2018; Esen et al., 2022; Meyer et al., 2019; Miller et al., 2018; Sharma & Honan, 2020). These results have been confirmed by other studies conducted on different age groups (Bozzato et al., 2021; Emvalotis & Koutsianou, 2018; Ferguson & Lezotte, 2020). Different instruments have been utilised to formulate an image of scientist based upon these findings. However, these instruments have been criticised because of their limitations (Table 1). These criticisms were addressed by the authors of this study in three categories: a) criticisms addressing the limitations of drawing skills, b) criticisms addressing the limitations of analysis and reporting, and c) the effect of prompting on drawings. The most commonly reported limitation of the use of DAST in the relevant literature concerns drawing skills. Sumrall (1995) stated that drawing on white paper and using a limited number of colours poses an obstacle to drawing skills. According to Rennie and Jarvis (1995), it is not possible to represent abstract ideas merely by drawing. In addition, Ferguson and Lezotte (2020) stated that the modern scientist image cannot be captured with simple drawing techniques, Christidou et al. (2021) argued that drawings would not be sufficient to reveal stereotypical images, and Walls (2022) concluded that drawings did not provide a rich view of the colours and races of scientists. Researchers reporting problems in the analysis and interpretation of DAST, on the other hand, stated that it is necessary to request additional information about the drawing from the person who drew it; thus, content analysis alone is not sufficient (Ball & Smith, 1992; Reinisch et al., 2017). In addition, researchers reported that if drawings have different meanings, the findings may be reported differently due to researcher bias (Thomas et al., 2006), and even if there are multiple drawings of scientists, only the first perceptions are captured by DAST (Farland-Smith, 2009).

Table 1. Instruments Used to Determine the Image of the Scientist and the Limitations of these Instruments.

Instrument	Limitations
DAST (Chambers, 1983)	• limitations of drawing skills (Christidou et al., 2021; Ferguson & Lezotte, 2020; Rennie & Jarvis, 1995; Sumrall, 1995; Walls, 2022), • limitations of analysis and reporting (Ball & Smith, 1992; Farland-Smith, 2009; Reinisch et al., 2017; Thomas et al., 2006), • the effect of prompting on drawings (Lamminpää et al., 2020; Laubach et al., 2012; Losh et al., 2008; Rapp & Kurby, 2008)
DAST-C Finson et al. (1995) and mDAST (Farland-Smith, 2012)	• rubric feature depending on the interpretation of the drawings (Edgerly et al., 2021; Lamminpää et al., 2020)
Interview (Mead & Metraux, 1957)	• difficulty in identifying student ideas (Chang et al., 2020) • analysis and interpretation difficulties (Christidou et al., 2021; Ferguson & Lezotte, 2020; Reinisch et al., 2017)
Word association tests (Ateş et al., 2021)	• very limited information (Alkış Küçükaydın, 2018), • difficulty in application, especially in younger age groups (Esen et al., 2022)
Metaphors (Dikmenli et al., 2012)	• difficulty in application, especially in younger age groups (Esen et al., 2022)

Researchers who thought that asking for extra information about DAST drawings from the participants was necessary also thought that the ideas were not fully reflected in the drawings and that an individual's drawing of a scientist implied the idea that there is actually a typical scientist for that individual (Lamminpää et al., 2020; Losh et al., 2008). According to Rapp and Kurby (2008), if there is a structure with more than one meaning in such drawings, the meaning cannot be revealed with DAST.

After these criticisms of the DAST, some researchers who prefer to draw the image of scientist have developed different interpretations for the use of the DAST. For example, Türkmen (2015) with Sharma and Honan (2020) asked their participants to add explanatory sentences while they were drawing. Laubach et al. (2012) suggested the use of interviews in addition to drawings. At the same time, Finson et al. (1995) developed the DAST-Checklist (DAST-C), and the modified DAST (mDAST) was developed by Farland-Smith (2012). Thus, it was reported that the drawings obtained from the DAST could be evaluated more easily. However, these developed tools have a rubric feature depending on the interpretation of the drawings. Accordingly, it was discussed whether the image of the scientist can be evaluated with closed-ended questions and test scoring (Edgerly et al., 2021; Lamminpää et al., 2020). These discussions paved the way for the development of a new Likert-type measurement tool.

The VoSAL

In the studies of the scientist image, different instruments were also included in addition to the DAST. In this context, interviews (Mead & Metraux, 1957), word association tests (Ateş et al., 2021), and metaphors (Dikmenli et al., 2012) were used. In addition, the proposal to quantitatively evaluate the views on the nature of science and the image of the scientist with Likert-type scales continued to be emphasised in a dominant way (Edgerly et al., 2021). The VoSAL is an instrument developed for this purpose. The VoSAL is a 5-point Likert-type instrument that encompasses the stereotypical appearance, locations, and scientific activity dimensions of scientists. The first use of the instrument was made with pre-service biology teachers, and it was found that the image of the scientist differed according to the grade level of the pre-service teachers (for more details, see Reinisch & Krell, 2023). However, these results obtained from the VoSAL, in which the image of the scientist was attempted to be determined for the first time using the self-report scale, have not yet been confirmed by different studies. Developed based on the RIASEC model (Wentorf et al., 2015), the VoSAL has a structure that can be used directly on pre-service teachers. For this reason, it was presented as an alternative to the DAST in studies to be carried out with pre-service teachers from different cultures and the results to be obtained were discussed in this context.

Theoretical underpinnings guiding DAST and VoSAL

In DAST, developed by Chambers (1983), participants are expected to reflect the image of a scientist in their minds onto a blank sheet of paper. This provides some advantages for younger age groups or disadvantaged groups with weak writing skills (Losh et al., 2008). Chambers (1983) reported that seven indicators (lab coat, eyeglasses, facial hair, symbols of research, symbols of knowledge, technology, and relevant captions) can be used in the analysis of these drawings. However, analyses of DAST plots are based on interpretative perspectives. This has necessitated more arguments to explain the drawings (Reinisch et al., 2017). In response, Farland-Smith (2012) encouraged scientists to develop different ideas about appearances, the locations where they work, and the activities they do. Wentorf et al. (2015), on the other hand, criticised DAST evaluations, stating that these approaches lead to stigmatising and prejudiced views about scientists. On the heels of that criticism, they developed the *Nature of Scientists Questionnaire*, based on Holland's (1963) RIASEC model. According to this model, there are typical features (realistic, investigative, artistic, social, enterprising, and conventional) that scientists may have. VoSAL makes reference to these typical features (Reinisch & Krell, 2023) and may be used to evaluate preservice teachers' concepts related to activities of scientists. Although the theoretical foundations that guide VoSAL and DAST are different, both instruments are aimed at determining images of scientists.

Purpose of the study

Although many studies have been carried out to determine the image of the scientist, the DAST has become quite established in image determination studies. However, it has been repeated many times in the literature that the DAST and drawing tools in general (Lamminpää et al., 2020) have some analysis and interpretation difficulties (Christidou et al., 2021; Ferguson & Lezotte, 2020; Reinisch et al., 2017). Accordingly, Chang et al. (2020) stated four reasons for using drawing as a measurement tool: a) its use as an alternative method due to the lack of development of writing skills, especially in younger age groups, b) revealing structures of thought that cannot be revealed by other methods, c) determining the characteristics of science subjects, and d) using it as a formative assessment tool to identify student ideas. However, these four reasons may not always be fulfilled in determining the image of the scientist. At this point, the VoSAL, developed by Reinisch and Krell (2023), draws attention. However, the VoSAL has a self-reporting structure. Therefore, its compatibility with the data obtained from the drawing tests has not been investigated yet. In addition, the VoSAL can be used as an alternative to the DAST, as it does not require drawing skills. There is a need to create a foresight for future research by considering the strengths and limitations of the VoSAL. In this context, a direct comparison of the VoSAL and DAST was made, thus it was discussed whether they can be used as alternatives to each other. Accordingly, the validity and reliability of the VoSAL was first tested on a Turkish sample. Researchers in Turkey have been working on images of scientists for a long time (Alkış Küçükaydın, 2018; Esen et al., 2022). The most basic tool used in these studies is DAST. However, all of the aforementioned limitations regarding the use of DAST also apply to Turkish studies. Therefore, a standardised measurement tool is needed as an alternative to DAST. Results from studies conducted with a valid and reliable measurement tool can provide valuable information for teachers, researchers, and policymakers. This understanding can be a guide for policymakers in particular in determining career choices related to science and identifying the effects of images of scientists on these choices. This is considered valuable in terms of adopting new approaches to the development of career perceptions related to science. There were two main aims to this research: firstly, to adapt VoSAL to Turkish and to conduct validity and reliability studies; and secondly, to compare the findings between VoSAL and DAST findings.

Study 1: adaptation of the VoSAL

Participants

The participants in this study consisted of undergraduate students studying at the faculty of education of a large university located in the central Anatolian Region of Turkey. The students were between the ages of 18–30 years (M = 21, *SD* = 1.61). A total of 625 undergraduate students from six different programmes (primary school teaching, preschool teaching, guidance and psychological counselling, Turkish teaching, primary school mathematics teaching, and science teaching) participated in the study. The students are receiving education at the university where the authors of this study work. Therefore, the convenience sampling technique was used in sample selection. In this context, the relevant instruments were administered to the students face-to-face on a voluntary basis during free lesson hours. Of the students, 476 were female and 149 were male. Moreover, 160 were studying in the 1st year, 179 in the 2nd year, 116 in the 3rd year, and 170 in the 4th year. The demographic characteristics of the students are presented in Table 2.

Translation procedures

In the study, first, permission was requested from the authors by e-mail for the adaptation of the VoSAL into Turkish and permission to use it. The authors shared the items in the VoSAL and the directions for use. Then, permission was obtained from the ethics committee of the university to which the researchers were affiliated, for the use of the instrument with the students of the faculty of education. After receiving permission, some principles were applied for the translation of the instrument (Beaton et al., 2000). First of all, the items in the VoSAL were reviewed and it was determined that there were no problems in terms of cultural conformity. Afterward, the language control steps were followed for the translation of the instrument items into Turkish, the back translation process, and the final version. The translation of the instrument items was done by the authors of this study. It was checked by an expert with a doctorate in the field of Turkish teaching to determine whether there were any errors in the translation

Table 2. Demographic Information of the Participants.

		Department						
		Guidance and Psychological Counseling	Turkish Teaching	Primary School Teaching	Elementary Mathematics Teaching	Pre-School Teaching	Science Teaching	Total
Gender	Female	102	110	78	113	54	19	476
	Male	28	36	20	40	9	16	149
Total		130	146	98	153	63	35	625
		Year 1	Year 2	Year 3	Year 4	Total		
Gender	Female	128	146	80	122	476		
	Male	32	33	36	48	149		
Total		160	179	116	170	625		

process. After the completion of the translation into Turkish, back-translation to the original language was carried out by a faculty member working in the translation and interpretation department. It was observed that there was no difference between the original items of the instrument and the translated items. After the instrument took its final form, a Turkish language expert checked the translation in terms of its intelligibility and suitability to the Turkish language structure, and the instrument was applied to 10 undergraduate students as a pilot study. The students in the pilot study were not involved in the actual practice. At the end of this process, the form in which the Turkish items of the VoSAL and the demographic characteristics of the participants were asked was converted into an online questionnaire and shared with undergraduate students.

Statistical analyzes used in the adaptation

After VoSAL was adapted to Turkish, the statistical analysis commenced. Construct validity was tested by looking at convergent validity and discriminant validity. Convergent validity assesses how well a construct is represented by its components. The average variance extracted (AVE) value and composite reliability (CR) value were taken into account in evaluating the composite validity (Hu & Bentler, 1999). In the evaluation of discriminant validity, the value for the criterion of maximum shared variance less than AVE was taken into account. All analyses were based on a $p < .05$ significance value.

Measures

Demographic questionnaire

In order to reach the demographic information of the participants in the study, they were asked about their age, gender, programme, and grade level.

VoSAL

The VoSAL is a 5-point Likert type (1 = not true … , 5 = totally true) instrument developed by Reinisch and Krell (2023). The instrument, which has a total of 29 items, contains items related to the stereotypical appearance (five items), inquiry location (four items), and scientific activity (20 items) of scientists. In this context, the instrument consists of three subscales. Therefore, each of the scales that make up the instrument is evaluated separately. The application time of the instrument varies between 5 and 10 min. Reinisch and Krell reported a Cronbach's alpha value of .70 for the stereotypical appearance and inquiry scales. The scientific activity scale, on the other hand, consists of subdimensions (realistic, investigative, artistic, teaching, conventional, and social), and the Cronbach's alpha values of these sub-dimensions range from .63 to .73. The values obtained within the scope of this study are presented in Table 3.

Data analysis

Within the scope of the study, first, the data set was reviewed in terms of extreme values. In order to determine the extreme values, Z-scores were evaluated and it was determined that the Z-scores of three participants were outside of the ±3 range. For this reason, the data of the three participants in question were removed from the data set. After removing the extreme values from the data set, normality tests were performed. In this context, the skewness and kurtosis coefficients were reviewed and it was seen that the values obtained (– .832 and .438 for appearance; – .293 and .606 for location, .524 and .592 for activity) were the limits of ±1. On the other hand, Büyüköztürk et al. (2011), the fact that the skewness and kurtosis coefficients were within the limit values of ±1 indicated that the data did not show a significant deviation from the normal distribution. After determining that the data were normally distributed, validity and reliability analyses were started.

Confirmatory factor analysis (CFA) was used to check the validity and reliability of the instrument on the Turkish sample. The maximum likelihood method was used for the CFA and the following values were adopted for the fit index: $\chi 2/sd < 3$; RMSEA, S-RMR < .08; GFI, AGFI, NFI, CFI > .90 (Browne & Cudeck, 1993; Hu & Bentler, 1999; Kelloway, 2015; Schermelleh-Engel et al., 2003). Convergence and reliability analyses (Cronbach's alpha and composite reliability) of each scale constituting the instrument were performed separately. SPSS 26.0 and AMOS 26.0 software were used for the data analysis.

Results

The CFA was applied in the validation study of the instrument on the Turkish sample. At this stage, analyses were repeated for each scale that made up the instrument. The first scale of the instrument concerns the appearance of the scientist. Excellent-fit-values were obtained without any modification in the CFA applied for the scale [$\chi 2$ (5, N = 625) = 1.243, RMSEA = .061, S-RMR = .031, GFI = .989, AGFI = .968, NFI = .985, CFI = .989]. The data for the single factor structure of the first scale are presented in Table 3.

According to the findings in Table 2, the scientist's appearance scale had a single factor structure and the factor loads of the items that make up the scale varied between .627 and .917. The construct validity of the measurement tool was tested with convergent validity, and its reliability was tested with Cronbach's alpha and CR coefficients. For convergent validity, AVE value greater than .50 was used (Fornell & Larcker, 1981). It was seen that the obtained value (.68) met this criterion. In the reliability analysis of the scale, it was seen that the Cronbach's alpha coefficient was .68 and the CR value was .96. Based on

Table 3. Values for the Appearance and Location Scale.

Item	Factor Loading	Cronbach's Alpha	AVE	CR
Appearance				
Item 1	.852	.68	.68	.96
Item 2	.817			
Item3	.627			
Item 4	.897			
Item 5	.917			
Location				
Item 1	.841	.90	.77	.97
Item 2	.940			
Item3	.902			
Item 4	.837			

the values obtained, it is possible to say that the scientist's appearance scale is valid and reliable for the Turkish sample (Gefen et al., 2000).

The CFA was performed again for the second scale to determine the perceptions about the location where the scientist works: $\chi 2$ (1, N = 625) = .619, RMSEA = .000, S-RMR = .003, GFI = .998, AGFI = .995, NFI = .998, CFI = .998. The validity and reliability values of the second scale were calculated and the obtained values are presented in Table 3.

According to Table 3, the AVE value (.77) calculated for the convergent validity of the location scale was greater than .50, and the Cronbach's alpha coefficient (.90) and CR value (.97) calculated for the reliability were quite high. This shows that the scale meets the criteria determined in terms of both convergent validity and reliability (Fornell & Larcker, 1981; Gefen et al., 2000; Hair et al., 1998).

The final scale of the VoSAL instrument relates to the activity of scientists. The activity scale consists of inquiry, realistic, social, conventional, artistic, and teaching dimensions. In the first trial, the 1st (the factor load = .410) and 2nd (the factor load = .480) items in the scale were excluded from the measurement model because their factor loads were below .50. Then, the analysis was repeated. Excellent fit values were obtained for the third scale, where no modifications were made [$\chi 2$ (119, N = 625) = .684, RMSEA = .060, S-RMR = .058, GFI = .909, AGFI = .906, NFI = .985, CFI = .909]. The last scale of the instrument has a six-factor structure and the values of the scale are presented in Table 4.

According to Table 4, the AVE value calculated for the convergent validity of the activity scale varied between .50 and .52 for the sub-dimensions. According to Fornell and Larcker's (1981) criteria, a factor load greater than .50 reports that convergent validity was achieved. However, the Cronbach's alpha value of each dimension that makes up the scale varied between .70 and .86. The CR value of the scale was .77 for the inquiry dimension, .76 for the realistic dimension, .89 for the social dimension, .84 for the conventional dimension, .84 for the artistic dimension, and .91 for the teaching dimension.

Table 4. Values for the Activity Scale.

Item	Factor Loading	Cronbach's Alpha	AVE	CR
Inquiry		.86	.53	.77
Item 3	.700			
Item 15	.728			
Item 19	.762			
Realistic				
Item 9	.759	.70	.52	.76
Item 12	.641			
Item 17	.760			
Social				
Item 4	.601	.86	.50	.89
Item 11	.602			
Item 14	.655			
Item 16	.681			
Item 20	.642			
Conventional				
Item 7	.808	.70	.52	.84
Item 10	.624			
Artistic				
Item 5	.763	.79	.51	.84
Item 8	.688			
Item 13	.696			
Teaching				
Item 6	.754	.73	.52	.91
Item 18	.700			

Therefore, it is possible to say that the internal consistency of the scale is high (Hair et al., 1998). The discriminant validity was also tested for the last scale consisting of six dimensions. The values obtained in this context are presented in Table 5.

Table 5 presents the relationship between the sub-dimensions that make up the activity scale and the square root of the AVE value. Accordingly, the square roots of the AVE values shown in bold in the Table 5 are greater than the other values in the same column and row. This shows that the scale has discriminant validity (Fornell & Larcker, 1981).

Discussions

In the first part of the study, the validity and reliability study of the VoSAL instrument developed by Reinisch and Krell (2023) was carried out by adapting it to Turkish culture. As in the original, it was observed that the VoSAL included three subscales in the Turkish sample. Excellent-fit-values were obtained in the CFA for the appearance scale. The obtained values revealed that the scientist's appearance scale was suitable for the Turkish sample (Gefen et al., 2000). As a result of the modification made between the 1st and 3rd items in the location scale, perfect fit values were achieved. Convergent validity and reliability analyses showed that the scale had a high level of adaptation to the Turkish sample (Fornell & Larcker, 1981; Hair et al., 1998). The activity scale consisted of six sub-dimensions in total: inquiry, realistic, social, conventional, artistic, and teaching. As a result of the first analysis, the factor loadings of the 1st and 2nd items were low, so they were excluded from the scale. This situation reduced the number of items in the conventional and teaching sub-dimensions of the scale. In multidimensional scales, the low number of items in the sub-dimensions may cause low reliability (Şencan, 2005). However, in the current study, both the Cronbach's alpha and CR values showed that the scale is reliable. It was determined that discriminant validity was provided for each of the scales (Fornell & Larcker, 1981). Therefore, the obtained values showed that the scale was compatible with the Turkish sample.

These values obtained on the Turkish sample can only be compared with the VoSAL because the literature on the determination of the scientist image with Likert-type self-report scales has not yet been formed. Based on this, it is possible to say that the Turkish VoSAL, which was adapted, is compatible with its original structure. Adaptation studies of VoSAL in different cultures are needed for deeper comparison studies. In addition, VoSAL does seem more satisfactory than other instruments utilised in the image of the scientist identification studies when it comes to validity. For example, Liang et al. (2008) developed a Likert-type 'Students' Understanding of Science and

Table 5. Discriminant Validity of the Activity Scale.

	1	2	3	4	5	6
Inquiry (1)	**.728**					
Realistic (2)	.702	**.721**				
Social (3)	.644	.649	**.707**			
Conventional (4)	.431	.602	.560	**.722**		
Artistic (5)	.706	.700	.706	.516	**.714**	
Teaching (6)	.617	.686	.703	.604	.640	**.722**

Scientific Inquiry' instrument, which deals with the elements of the nature of science, but does not examine the image of the scientist separately. Similarly, Sjøberg and Schreiner (2019) used DAST in the relevance of science education projects by adding it to survey items outside the drawing instrument. However, in these applications, the analysis of the image of the scientist has not been dealt with separately. This shows that VoSAL is currently an adequate instrument that can be used in the image of scientist studies.

Study 2: comparison of the DAST and VoSAL

Participants

This part of the study, it was aimed to compare the data obtained from the VoSAL and DAST validated on the Turkish sample. The participants of the study consisted of undergraduate students studying at the faculty of education of a large university located in the central Anatolian Region of Turkey and who were not included in the first study. Therefore, the participants in Study-2 are not a subset of the 625 participants in Study-1. The students are between the ages of 18 and 23 years ($M = 20$, $SD = 1.05$). A total of 46 undergraduate students from five different programmes (primary school teaching, preschool teaching, science teaching, Turkish teaching, and primary school mathematics teaching) participated in the study. Of the students, 36 were female and 10 were male.

Procedures

In this part of the study, the participants were instructed to draw a scientist by distributing blank A4 paper and coloured pencils. The participants were asked to write their age, gender, and programme on the back of the same paper. The VoSAL, whose validity and reliability were tested on the Turkish sample, was applied to the participants who finished the drawing. The participants were given a total of 40 min to draw and respond to the VoSAL instrument. In order to protect the anonymity of the participants, they were asked to write their own nicknames or ID number on the DAST and VoSAL forms.

Data analysis

Within the scope of the study, the data of the DAST and VoSAL were analyzed separately. In this context, first, the DAST data were evaluated. The mDAST, developed by Farland-Smith (2012), was used in the evaluation of the DAST. There were different approaches used in the literature to evaluate the DAST data (Christidou et al., 2012; Finson et al., 1995; Meyer et al., 2019). However, the mDAST was preferred so the data obtained from the DAST and VoSAL could be analyzed more easily under categories (appearance, location, and activity). In the mDAST, drawings are evaluated separately in the categories of appearance, location, and activity. One point given to the drawings made under these categories means 'sensationalized', 2 points 'traditional', 3 points 'broader than traditional', and 0 points 'can't be categorized'. The higher the values obtained by scoring the plots, the less stereotypical the representations are. The drawings made by the participants were evaluated independently by the authors of this study. Therefore, after the 46 drawings were coded according to the mDAST, reliability was checked for both the

Cohen kappa (κ) and Krippendorff's alpha coefficient (α) values. The 'SPSS syntax (mkappasc.sps)' file was used to calculate the κ statistic, and the 'SPSS syntax (kalpha.sps)' file was used for α. Accordingly, for appearance κ = 0.83; α = 0.80, for location κ = 0.82; α = 0.81, and for activity κ = 0.81; α = 0.83 values showed high coder agreement (Krippendorff, 1995; McHugh, 2012).

In the evaluation of the answers to the questions in the VoSAL, the intervals were assumed to be equal and the score interval for the arithmetic averages was calculated as .80 [Score Interval = (Highest Value – Lowest Value)/5 = (5–4)/5 = 4/5 = 0.80]. According to this calculation, 1.00–1.80 = not false, 1.81–2.60 = false, 2.61–3.40 = neither false nor true, 3.41–4.20 = true, and 4.21–5.00 = totally true. Microsoft Excel was used to analyze the data collected through VoSAL..

Results

Findings related to appearance

When the mDAST indicators were examined, it was seen that 57% of the participants had the image of a 'traditional' scientist and 39% had the image of a 'broader than traditional' scientist. Moreover, 4% of the drawings of the participants 'cannot be categorized'. However, no drawing that could be coded as 'sensationalized' was found under this category. The average score from the mDAST was 2.30 (*SD* = .69), which indicated a traditional image of the scientist. An example drawing made under this category is presented in Figure 1a. An example drawing on the appearance of the scientist from a 'broader than traditional' perspective is presented in Figure 1b, and an example of a 'cannot be categorized' drawing is presented in Figure 1c.

It was seen that the mean score of the VoSAL in the category of the scientist's appearance was 3.56 (*SD* = .62). For the scale scored in a 5-point Likert type, this value was close to the expression 'true'. The VoSAL scale asks for the age of scientists in general and the clothes they prefer to wear at work. Therefore, the participants generally imagine the scientist as above a certain age (older than 50 years) and think that they wear protective clothing at work. Under this category, it was seen that the responses obtained from both the mDAST and the appearance scale of the VoSAL confirmed each other and both represented the traditional understanding.

Findings related to location

The drawings made by the participants continued to be examined in terms of the mDAST indicators, and the drawings pointing to the place where the scientist worked were evaluated. As a result of the related evaluation, it was seen that 67% of the participants made drawings reflecting the 'traditional' images and 9% reflected 'broader than traditional' images. No drawing that could be coded as 'sensationalized' was found. However, 24% of the drawings 'cannot be categorized' because in 24% of the drawings, an image pointing to the place where the scientist was working was not encountered. In their drawings, the participants focused more on the appearance of the scientist and his/her work but did not make detailed drawings about the place where he/she worked. The average score obtained from the mDAST in the location category was 1.60 (*SD* = .95). This value

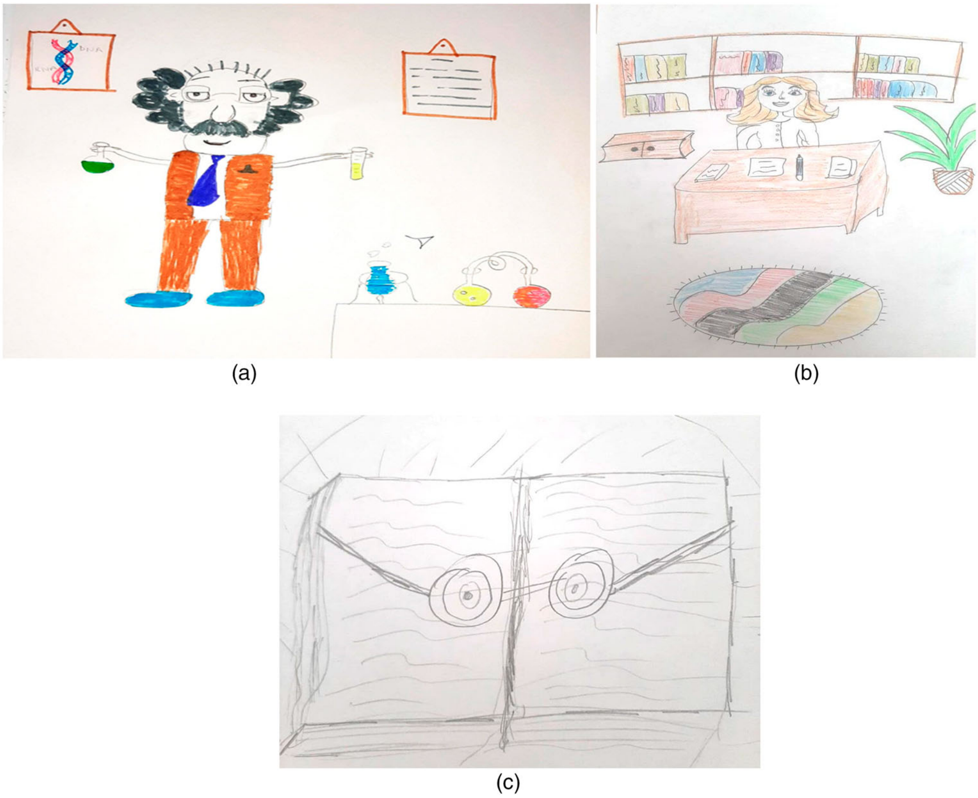

Figure 1. (a) Example of a traditional drawing of the appearance of the scientist. (b) Example of a broader than traditional drawing of the appearance of the scientist. (c) Example of a cannot be categorized drawing of the appearance of the scientist.

indicated ‘traditional’ drawings at an average level. An example drawing reflecting this category is presented in Figure 2a. An example of a ‘broader than traditional’ drawing regarding the location of the scientist is presented in Figure 2b, and a drawing example of ‘cannot be categorized’ is presented in Figure 2c.

Regarding the location where the scientist works, the average score obtained from the location scale of the VoSAL instrument was 3.73 (*SD* = .72). For the scale scored in a 5-point Likert type, this value was close to the expression ‘true’. That is, according to the participants, scientists are located in work areas that can be classified as traditional (for example, the laboratory). This showed that participants in the location scale of both the mDAST and VoSAL had similar results regarding the scientist’s workplace.

Findings related to activity

When the drawings of the participants about the scientist’s activity were examined in terms of the mDAST indicators, it was seen that 72% of the participants had a ‘traditional’ image, 11% had a ‘broader than traditional’ image, and 4% had a ‘sensationalized’ image of a scientist. According to the drawings obtained, 13% ‘cannot be categorized’. Under this category, the participants generally drew activity drawings

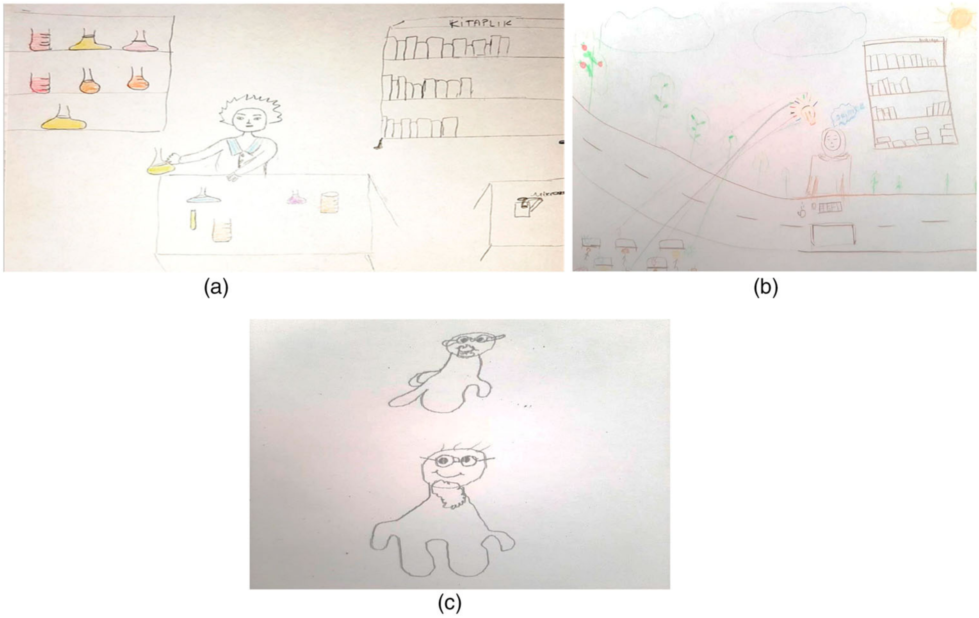

Figure 2. (a) Example of a traditional drawing about the location where the scientist works. (b) Example of a broader than traditional drawing about the location where the scientist works. (c) Example of a cannot be categorized drawing about the location where the scientist works

representing experiments, experimental animals, or a laboratory environment (Figure 3a). The average score obtained from the mDAST in the activity category was 1.80 (*SD* = .80). This value indicated 'traditional' drawings at an average level. When it comes to the scientist's activity, an example drawing of the 'broader than traditional' category is provided in Figure 3b, an example of 'cannot be categorized' is presented in Figure 3c., and an example of 'sensationalized" is presented in Figure 3d.

Regarding the scientist's activity, the average score obtained from the activity scale of the VoSAL instrument was 3.95 (*SD* = .69). The score equivalent of this value at the item level is the expression 'true'. This showed that the participants generally think of scientists in a structure that is characterised as traditional (unrealistic discoveries or constantly working on experimental animals). These findings obtained from both the mDAST and VoSAL scale supported each other. The responses of the participants to each category of the mDAST and subscales of the VOSAL are presented in Appendix in detail. Accordingly, the average values obtained from the instruments were close to each other.

Discussion

In this study, the data of the DAST and VoSAL were evaluated in the categories of appearance, location, and activity, and the drawings of the DAST were analyzed using the mDAST (Farland-Smith, 2012). The data obtained from both the mDAST and VoSAL regarding the appearance of scientists showed that pre-service teachers had a 'traditional' image of scientists. Although it has been observed that the traditional image of scientists has gradually changed in recent studies (McCarthy, 2015; Miller et al., 2018), it is understood that most stereotypical features remain the same (Bozzato et al., 2021; Esen et al., 2022; Lamminpää et al., 2020; Reinisch et al., 2017). This situation persisted in the present study as well. However, the sensationalised image structure of Mead and Metraux (1957), which made dangerous experiments and was described as crazy, was not included in the drawings. The VoSAL, on the other hand, does not contain an item that will reveal this image. In addition, there is no item in the VoSAL that questions the gender of scientists. However, the content that is traditionally categorised in the mDAST takes into account whether scientists are male or female. This shows that if the VoSAL alone is used, only the age and clothing of the scientists can be obtained under this category. However, the expression 'scientist' used in both the DAST and VoSAL can be directive (Christidou et al., 2012). This orientation may have had an effect on the drawings. Bernard and Dudek (2017) developed the Indirect Draw-a-Scientist Test (InDAST) in order to free the participants from this orientation and reported that it disproved the theories about the gender of the scientist. Therefore, it is possible that orientation was dominant in the drawings obtained from the mDAST regarding the appearance of the scientist. However, it is not possible to say that the questions covered by the VoSAL were sufficient. In this context, the results to be obtained in comparisons to be made with the rubrics that offer different interpretations of the VoSAL and DAST may be different.

When the mDAST and VoSAL data were examined, it was seen that the findings regarding the location where the scientist worked pointed to the traditional image structure. This conclusion was also reached in previous studies (Medina-Jerez et al., 2011; Subramaniam et al., 2013). In addition, in the study of Gheith and Aljaberi (2019), pre-service teachers mostly depicted scientists in closed environments. Similarly, in the study of Christidou et al. (2021), teachers imagined scientists in the laboratory. Therefore, it is possible to say that the traditional image structure of the location where the scientist works still remains the same. There was no drawing that could be described as sensationalised in the category of the place where the scientist worked, but 24% of the drawings could not be categorised. This is a common methodological issue with the DAST (Avraamidou, 2013; Lamminpää et al., 2020; Losh et al., 2008). Ateş et al. (2021) reported that abilities come to the forefront in the DAST drawings and this affects the analysis. Therefore, the inability to associate the drawings with a certain category in this study was a serious limitation for the DAST. When the responses to the VoSAL were analyzed, another limitation was encountered because there are four items under this category in the VoSAL. However, in the study of Lamminpää et al. (2020), different drawings of scientists in the classroom began to be observed. Reinisch and Krell (2023) stated that they added additional items for the location scale to the related instrument during the development phase of the VoSAL, but they deleted the items due to the low factor load in the first analysis. Therefore, if these and similar items (for example, outdoor locations) are added, the answers by the participants may vary. However, it is also clear that the current structure of the VoSAL offers an insight into drawings that cannot be categorised with the mDAST. However, if the VoSAL and DAST are applied together, ideas that are not included in the VoSAL can be revealed with the DAST. This situation indicates that the VoSAL and DAST should not be used as alternative measurement tools instead of each other, but rather it would be healthier to consider them in a complementary structure.

It was seen that the findings of the instruments regarding the activity category supported each other and that the pre-service teachers had a traditional image structure of the activities of scientists. Therefore, the data obtained from the mDAST and VoSAL confirmed each other for the activity category. However, both instruments had limitations at this stage. First, ideas remain within certain limits according to the VoSAL because the VoSAL instrument is the scientist's activity that divides it into sub-dimensions as inquiry, realistic, social, conventional, artistic, and teaching. However, there may be ideas that do not belong to these dimensions. For example, Bozzato et al. (2021) conducted a study on Italian schoolchildren and evaluated their drawings with the mDAST and DAST-C, and reported that the image structures related to the activity of the scientist differed (studying, working on the computer, conducting an experiment, etc.). This showed that some image structures are not included in the VoSAL. However, 13% of the drawings could not be categorised in the activity category of the mDAST. This may have been related to the DAST's own directive, as well as the fact that details are not always included in the drawings because in the DAST's directive, participants are instructed to draw a scientist. In this case, the participants may be inclined to only draw the external appearance of the scientist and may need additional instruction to draw the internal features, such as the place where they work or their activities. Lamminpää et al. (2020), in their study using 'Draw a comic about how you think science is made', asked students to create comics and obtained rich data about the activities of scientists. In these drawings, it was seen that scientists were depicted as individuals who think, evaluate, experiment, discuss results, and participate in scientific conferences. Therefore, there are some limitations in determining the image structures of scientists' activities in terms of both the mDAST and VoSAL. From this point of view, the idea of choosing either the DAST or the VoSAL over the other may not be realistic.

Although the DAST has proven to be a versatile tool for research on scientists and science concepts, it may be insufficient to determine the perceptions of people with multiple images of scientists (Lamminpää et al., 2020). In addition, the stereotypical images in the drawings shed light on only some of the thoughts and perceptions (Christidou et al., 2021). In this sense, the VoSAL can offer a broader framework for presenting the image of the scientist. The questions asked about the appearance, location, and activity of the scientists seem quite clear and explanatory. However, focusing only on certain dimensions and questions in the VoSAL may lead to the evaluation of students' images of scientists only in terms of the questions asked. This shows that the DAST and VoSAL have their own strengths and weaknesses. Although the data obtained from the study showed that the data obtained from the mDAST confirmed the VoSAL, it is clear that both instruments have shortcomings. Reinisch et al. (2017) reported that one-third of pre-service teachers cannot clearly describe their ideas in drawings. This indicates that additional instruction should be submitted to the drawing. However, the VoSAL, which has only self-report quality, includes standardised statements. Apart from these expressions, it does not seem possible to reveal the image structure of individuals. As briefly stated by Reinisch and Krell (2023), the VoSAL can be used as an alternative tool to the DAST. However, considering the standardised nature of the VoSAL and the limitations of the DAST due to methodological issues, it may be possible to obtain richer data through the combined use of these instruments.

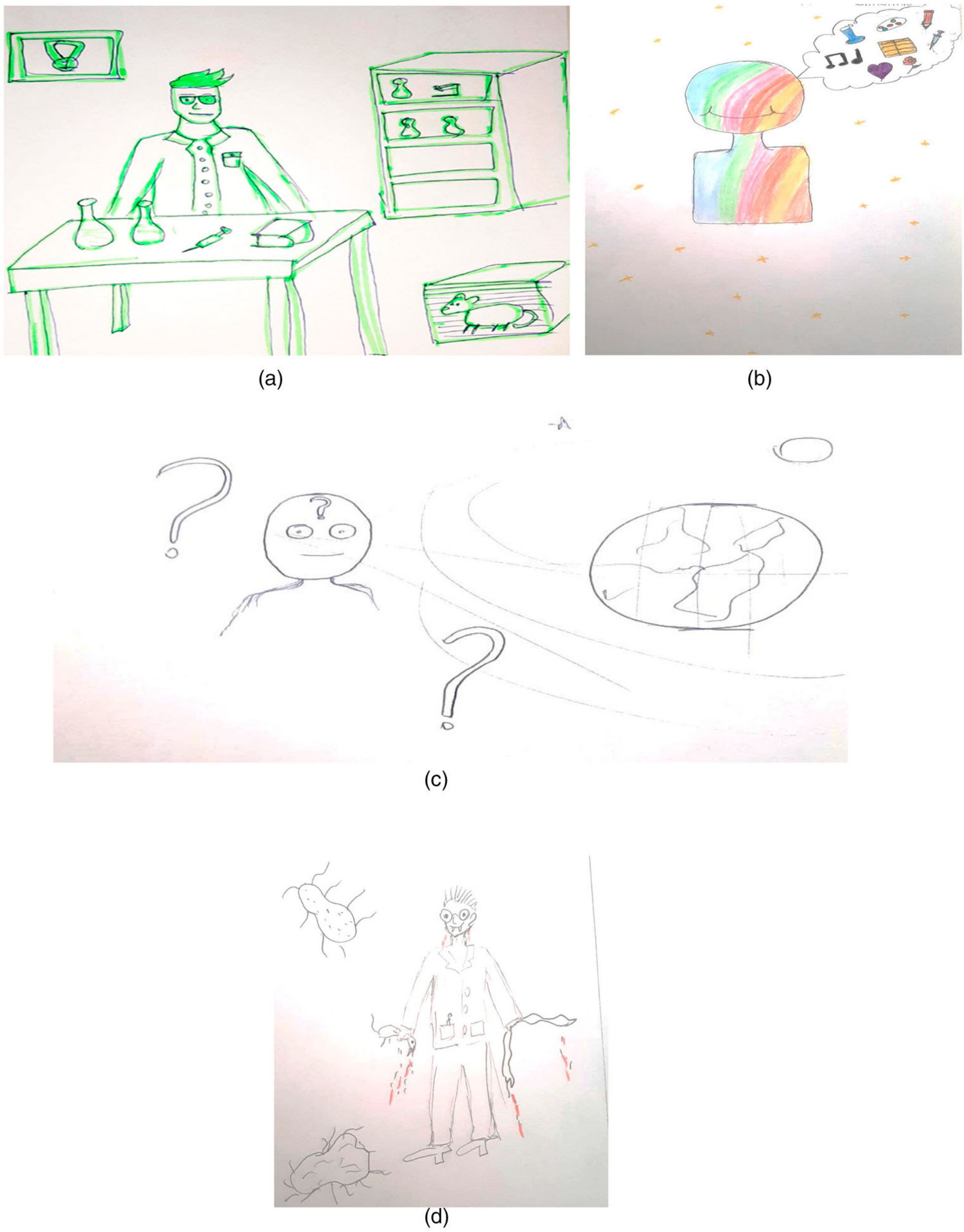

Figure 3. (a) Example of a traditional drawing of the scientist's activity. (b) Example of a broader than traditional drawing of the scientist's activity. (c) Example of a cannot be categorized drawing of the scientist's activity. (d) Example of sensationalized drawing of the scientist's activity.

Limitations and recommendations

In this study, the validity and reliability study of the VoSAL, in which the image of the scientist is attempted to be determined through self-report, was conducted in a Turkish sample, and then the data obtained from the VoSAL and DAST were compared. However, there were some limitations specific to the study in this context. First, the

VoSAL (Reinisch & Krell, 2023) and mDAST (Farland-Smith, 2012) scoring were compared in the study. The DAST-C form is also widely used for scoring the DAST (Finson et al., 1995). In this form, different external appearance features of the scientist, such as messy hair and glasses, can also be scored. In addition, different drawing tools have recently been developed that emphasise examining the emotions of scientists (Christidou et al., 2021). Additionally, in the InDAST (Bernard & Dudek, 2017), drawings are made without using the expression of a scientist. Using these drawing tools, a wider variety of scientist image detection tools can be compared in future research. Another limitation of the study was related to the use of the DAST. In the study, the pre-service teachers were asked to draw a scientist and the drawings were evaluated according to the intention of the participants. In the related literature, in addition to drawings, participants are also asked to write sentences describing their drawings (Emvalotis & Koutsianou, 2018; Sharma & Honan, 2020; Türkmen, 2015). Therefore, researchers who will conduct studies on this subject may ask students to write explanatory sentences or use mixed study designs to be sure of their intentions. Finally, the VoSAL is still a very young instrument, and this was the first time it was compared to the long-established DAST. Based on the fact that the image of the scientist is influenced by culture (Farland-Smith, 2009; Ferguson & Lezotte, 2020; Miller et al., 2018; Narayan et al., 2013), the study can be repeated in different cultures. Thus, the cultural context of the comparison of the VoSAL and DAST will also be addressed.

Conclusions

In the current study, the validity and reliability study of the VoSAL in a Turkish sample was conducted with 625 pre-service teachers from different departments. The findings showed that the VoSAL consisted of three subscales as in its original structure. Accordingly, the first scale is related to the appearance of the scientist and consists of a single factor. The second scale is related to the location where the scientist works and this scale consists of a single factor. The third scale, which is related to the activity of the scientist, consists of a six-factor structure. In addition, the VoSAL and DAST data were compared on a different sample in the study. The obtained results showed that the findings obtained by scoring the VoSAL and DAST confirmed each other. However, considering the advantages offered separately by the DAST and VoSAL and the limitations of these instruments, it is predicted that it will be healthier to use the relevant instruments together rather than interchangeably. This study employed DAST and VoSAL to collect data with the help of pre-service teachers from different branches. The variety of backgrounds represented in the research shows that VoSAL can be an effective instrument in determining the image of scientists within different groups. This will certainly help in pinning down the variables that define the image of the scientist according to pre-service teachers. In addition, it can be used to form the theoretical framework of causal studies by taking advantage of Likert-type tools. It can certainly also be used in modelling studies related to the factors affecting the image of the scientist.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Ethical approval

The study was in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Ethical approval for the research written permission was also granted by Selçuk University Social and Human Sciences Scientific Research Ethics Committee in meeting number 22 on 16/05/2022.

Informed consent

Informed consent was obtained from all individual.

ORCID

Menşure Alkış Küçükaydın http://orcid.org/0000-0003-4410-1279
Seher Esen http://orcid.org/0000-0002-3569-1185

Data availability

The datasets created during and/or analysed during the current study are available from the corresponding author on reasonable request.

References

Alkış Küçükaydın, M. (2018). An action research on the scientist image of 4th grade students. *Asia-Pacific Forum on Science Learning and Teaching*, *19*(1), 1–22.

Ateş, Ö, Ateş, A. M., & Aladağ, Y. (2021). Perceptions of students and teachers participating in a science festival regarding science and scientists. *Research in Science & Technological Education*, *39*(1), 109–130. https://doi.org/10.1080/02635143.2020.1740666

Avraamidou, L. (2013). Superheroes and supervillains: Reconstructing the mad-scientist stereotype in school science. *Research in Science & Technological Education*, *31*(1), 90–115. https://doi.org/10.1080/02635143.2012.761605

Ball, M. S., & Smith, G. W. H. (1992). *Analyzing visual data*. Sage.

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, *25*(24), 3186–3191. https://doi.org/10.1097/00007632-200012150-00014

Bernard, P., & Dudek, K. (2017). Revisiting students' perceptions of research scientists –outcomes of an indirect draw-a-scientist test (InDAST). *Journal of Baltic Science Education*, *16*(4), 562–575. https://doi.org/10.33225/jbse/17.16.562

Bianchini, J. A., Cavazos, L. M., & Helms, J. V. (2000). From professional lives to inclusive practice: Science teachers and scientists' views of gender and ethnicity in science education. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *37*(6), 511–547.

Bozzato, P., Fabris, M. F., & Longobardi, C. (2021). Gender, stereotypes and grade level in the draw-a-scientist test in Italian schoolchildren. *International Journal of Science Education*, *43*(16), 2640–2662. https://doi.org/10.1080/09500693.2021.1982062

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen, & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). SAGE.

Büyüköztürk, B., Çokluk, Ö, & Köklü, N. (2011). *Statistics for social sciences* (7th ed.). Pegem Academy.

Chambers, D. W. (1983). Stereotypic images of the scientist: The draw- a scientist test. *Science Education*, *67*(2), 255–265. https://doi.org/10.1002/sce.3730670213

Chang, H. Y., Lin, T. J., Lee, M. H., Lee, S. W. Y., Lin, T. C., Tan, A. L., & Tsai, C. C. (2020). A systematic review of trends and findings in research employing drawing assessment in science education. *Studies in Science Education*, *56*(1), 77–110. https://doi.org/10.1080/03057267.2020.1735822

Christidou, V., Bonoti, F., & Hatzinikita, V. (2021). Drawing a scientist: Using the emo-DAST to explore emotional aspects of children's images of scientists. *Research in Science & Technological Education*, *Ahead-of-Print*, 1–22. https://doi.org/10.1080/02635143.2021.1998770

Christidou, V., Bonoti, F., & Kontopoulou, A. (2016). American and Greek children's visual images of scientists. *Science & Education*, *25*(5), 497–522. https://doi.org/10.1007/s11191-016-9832-8

Christidou, V., Hatzinikita, V., & Samaras, G. (2012). The image of scientific researchers and their activity in Greek adolescents' drawings. *Public Understanding of Science*, *21*(5), 626–647. https://doi.org/10.1177/0963662510383101

DeWitt, J., Archer, L., & Osborne, J. (2014). Science-related aspirations across the primary–secondary divide: Evidence from two surveys in England. *International Journal of Science Education*, *36*(10), 1609–1629. https://doi.org/10.1080/09500693.2013.871659

Dikmenli, M., Çardak, O., & Yener, D. (2012). Science student teachers' metaphors for scientists. *Energy Education Science and Technology Part B: Social and Educational Studies*, *4*(1), 51–66.

Edgerly, H. S., Kruse, J. W., & Wilcox, J. L. (2021). Quantitatively investigating in-service elementary teachers' nature of science views. *Research in Science Education, 52*, 1467–148. http://doi.org/10.1007/s11165-021-09993-7

Emvalotis, A., & Koutsianou, A. (2018). Greek primary school students' images of scientists and their work: Has anything changed? *Research in Science & Technological Education*, *36*(1), 69–85. https://doi.org/10.1080/02635143.2017.1366899

Esen, S., Türkyılmaz, S., & Alkış Küçükaydın, M. (2022). Examining the effect of scientist biographies prepared by digital storytelling on primary school students' image of the scientist. *Pamukkale University Journal of Education*, *55*, 155–179. https://doi.org/10.9779/pauefd.1003461

Farland-Smith, D. (2009). How does culture shape students' perceptions of scientists? Cross-national comparative study of American and Chinese elementary students. *Journal of Elementary Science Education*, *21*(4), 23–42. https://doi.org/10.1007/BF03182355

Farland-Smith, D. (2012). Development and field test of the modified draw-a-scientist test and the draw-a-scientist rubric. *School Science and Mathematics*, *112*(2), 109–116. https://doi.org/10.1111/j.1949-8594.2011.00124.x

Ferguson, S. L., & Lezotte, S. M. (2020). Exploring the state of science stereotypes: Systematic review and meta-analysis of the draw-a-scientist checklist. *School Science and Mathematics*, *120*(1), 55–65. https://doi.org/10.1111/ssm.12382

Finson, K. D., Beaver, J. V., & Cramond, V. L. (1995). Development and field test of a checklist for the draw-a-scientist test. *School Science and Mathematics*, *95*(4), 195–205. https://doi.org/10.1111/j.1949-8594.1995.tb15762.x

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50. https://doi.org/10.1177/002224378101800104

Gefen, D., Straub, D., & Boudreau, M. C. (2000). Structural equation modeling and regression: Guidelines for research practice. *Communications of the Association for Information Systems*, *4*(1), 7–11. https://doi.org/10.17705/1CAIS.00407

Gheith, E., & Aljaberi, N. M. (2019). The image of scientists among pre-service classroom and child education teachers in Jordan. *International Journal of Instruction*, *12*(4), 561–578. https://doi.org/10.29333/iji.2019.12436a

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis* (5th ed.). Prentice Hall.

Holland, J. L. (1963). Explorations of a theory of vocational choice and achievement. *Psychological Reports*, *12*(2), 547–594. https://doi.org/10.2466/pr0.1963.12.2.547

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Kelloway, E. K. (2015). *Using Mplus for structural equation modeling* (2nd ed). SAGE

Krippendorff, K. (1995). On the reliability of unitizing continuous data. *Sociological Methodology*, *25*, 47–76. https://doi.org/10.2307/271061

Lamminpää, J., Vesterinen, V. M., & Puutio, K. (2020). Draw-a-science-comic: Exploring children's conceptions by drawing a comic about science. *Research in Science & Technological Education*, 41(1), 39–60. https://doi.org/10.1080/02635143.2020.1839405

Laubach, T. A., Crofford, G. D., & Marek, E. A. (2012). Exploring native American students' perceptions of scientists. *International Journal of Science Education*, *34*(11), 1769–1794. https://doi.org/10.1080/09500693.2012.689434

Liang, L. L., Chen, S., Chen, X., Kaya, O. N., Adams, A. D., Macklin, M., & Ebenezer, J. (2008). Assessing preservice elementary teachers' views on the nature of scientific knowledge: A dual-response instrument. *Asia-Pacific Forum on Science Learning and Teaching*, *9*(1), 1–20.

Losh, S. C., Wilke, R., & Pop, M. (2008). Some methodological issues with "draw a scientist tests" among young children. *International Journal of Science Education*, *30*(6), 773–792. https://doi.org/10.1080/09500690701250452

McCann, F. F., & Marek, E. A. (2016). Achieving diversity in STEM: The role of drawing-based instruments. *Creative Education*, *07*(15), 2293–2304. https://doi.org/10.4236/ce.2016.715223

McCarthy, D. (2015). Teacher candidates' perceptions of scientists: Images and attributes. *Educational Review*, *67*(4), 389–413. https://doi.org/10.1080/00131911.2014.974510

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276–282. https://doi.org/10.11613/BM.2012.031

Mead, M., & Metraux, R. (1957). The image of the scientist among high school students: A pilot study. *Science*, *126*(3270), 384–390. https://doi.org/10.1126/science.126.3270.384

Medina-Jerez, W., Middleton, K. V., & Orihuela-Rabaza, W. (2011). Using the DAST-C to explore columbian and bolivian students' images of scientists. *International Journal of Science and Mathematics Education*, *9*(3), 657–690. https://doi.org/10.1007/s10763-010-9218-3

Meyer, C., Guenther, L., & Joubert, M. (2019). The draw-a-scientist test in an African context: Comparing students' (stereotypical) images of scientists across university faculties. *Research in Science & Technological Education*, *37*(1), 1–14. https://doi.org/10.1080/02635143.2018.1447455

Milford, T. M., & Tippett, C. D. (2013). Preservice teachers' images of scientists: Do prior science experiences make a difference? *Journal of Science Teacher Education*, *24*(4), 745–762. https://doi.org/10.1007/s10972-012-9304-1

Miller, D. I., Nolla, K. M., Eagly, A. H., & Uttal, D. H. (2018). The development of children's gender-science stereotypes: A meta-analysis of 5 decades of US draw-a-scientist studies. *Child Development*, *89*(6), 1943–1955. https://doi.org/10.1111/cdev.13039

Narayan, R., Park, S., Peker, D., & Suh, J. (2013). Students' images of scientists and doing science: An international comparison study. *Eurasia Journal of Mathematics, Science & Technology Education*, *9*(2), 115–129. https://doi.org/10.12973/eurasia.2013.923a

Rapp, D. N., & Kurby, C. A. (2008). The 'ins' and 'outs' of learning: Internal representations and external visualizations. In J. Gilbert, M. Reiner, & M. B. Nakhleh (Eds.), *Visualization: Theory and practice in science education* (pp. 29–52). Springer.

Reinisch, B., & Krell, M. (2023). Assessing pre-service teachers' views of scientists, their activities and locations: The VoSAL instrument. *Research in Science Education*, *53*, 139–153. https://doi.org/10.1007/s11165-022-10046-w

Reinisch, B., Krell, M., Hergert, S., Gogolin, S., & Krüger, D. (2017). Methodical challenges concerning the draw-A-scientist test: A critical view about the assessment and evaluation of learners' conceptions of scientists. *International Journal of Science Education*, *39*(14), 1952–1975. https://doi.org/10.1080/09500693.2017.1362712

Rennie, L. J., & Jarvis, T. (1995). Children's choice of drawings to communicate their ideas about technology. *Research in Science Education*, *25*(3), 239–252. https://doi.org/10.1007/BF02357399

Rosenthal, D. B. (1993). Images of scientists: A comparison of biology and liberal studies majors. *School Science and Mathematics*, *93*(4), 212–216. https://doi.org/10.1111/j.1949-8594.1993.tb12227.x

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, *8*(2), 23–74.

Şencan, H. (2005). *Reliability and validity in social and behavioral measurements*. Seçkin Publish.

Sharma, R. A., & Honan, E. (2020). Fijian pre-service teachers' ideas about science and scientists. *Journal of Science Teacher Education*, *31*(3), 335–357. https://doi.org/10.1080/1046560X.2019.1706904

Sjøberg, S., & Schreiner, C. (2019). ROSE (The Relevance of Science Education). The development, key findings and impacts of an international low cost comparative project. Final Report, Part 1 (of 2). https://www.researchgate.net/publication/335664683_ROSE_The_Relevance_of_Science_Education_The_development_key_findings_and_impacts_of_an_international_low_cost_comparative_project_Final_Report_Part_1_of_2

Subramaniam, K., Esprívalo Harrell, P., & Wojnowski, D. (2013). Analyzing prospective teachers' images of scientists using positive, negative and stereotypical images of scientists. *Research in Science & Technological Education*, *31*(1), 66–89. https://doi.org/10.1080/02635143.2012.742883

Sumrall, W. J. (1995). Reasons for the perceived images of scientists by race and gender of students in grades 1–7. *School Science and Mathematics*, *95*(2), 83–90. https://doi.org/10.1111/j.1949-8594.1995.tb15733.x

Tan, A. L., Jocz, J. A., & Zhai, J. (2017). Spiderman and science: How students' perceptions of scientists are shaped by popular media. *Public Understanding of Science*, *26*(5), 520–530. https://doi.org/10.1177/0963662515615086

Thomas, M. D., Henley, T. B., & Snell, C. M. (2006). The draw a scientist test: A different population and a somewhat different story. *College Student Journal*, *40*(1), 140–149.

Türkmen, H. (2015). Still persistent global problem of scientists' image. *Asia-Pacific Forum on Science Learning and Teaching*, *16*(1), 1–21.

Walls, L. (2022). A critical race theory analysis of the draw-a-scientist test: Are they really that white? *Cultural Studies of Science Education*, *17*(1), 141–168. https://doi.org/10.1007/s11422-022-10107-6

Wentorf, W., Höffler, T. N., & Parchmann, I. (2015). Schülerkonzepte über das Tätigkeitsspektrum von Naturwissenschaftlerinnen und Naturwissenschaftlern: Vorstellungen, korrespondierende Interessen und Selbstwirksamkeitserwartungen[Student concepts about the range of activities of natural scientists: Ideas, corresponding interests and expectations of self-efficacy]. *Zeitschrift für Didaktik der Naturwissenschaften*, *21*(1), 207–222. https://doi.org/10.1007/s40573-015-0035-7

Appendix 1

Descriptive analysis results of the mDAST and VoSAL.

mDAST (n = 46)		*VoSAL* (n = 46)				
Appearance		*Appearance*	*M*	*SD*	*Min*	*Max*
		The typical natural scientist …				
Cannot be categorized	2 (% 4.37)	… wears protective cloth at work.	3.93	.95	1.00	5.00
Sensationalized	–	… is rather old (older than 50 years).	2.47	1.11	1.00	5.00
Traditional	26 (% 56.52)	… wears everyday clothes at work. *	3.43	1.06	1.00	5.00
Broader than traditional	18 (% 39.13)	… wears a lab coat at work.	4.08	.78	1.00	5.00
		… wears safety glasses at work.	3.86	.93	1.00	5.00
Location		***Location***				
		The typical natural scientist works …				
Cannot be categorized	11 (% 23.91)	… in the lab.	4.06	.85	2.00	5.00
Sensationalized	–	… in the nature.	3.69	.96	1.00	5.00
Traditional	31 (% 67.39)	… outside.	3.56	1.00	1.00	5.00
Broader than traditional	4 (% 8.69)	… in the open country.	3.63	1.10	1.00	5.00
Activity		***Activity***				
		The typical natural scientist performs the following activities regularly:				
Cannot be categorized	6 (% 13.04)	*Inquiry*				
Sensationalized	2 (% 4.37)	Analyze and interpret results from experiments.	4.36	.79	2.00	5.00
Traditional	33 (71.73)	Recognize relationships in measured data.	4.10	.87	1.00	5.00
Broader than traditional	5 (%10.86)	Develop ideas for new research approaches.	4.26	.80	1.00	5.00
		Realistic				
		Carry out an investigation.	4.26	.74	1.00	5.00
		Make a protocol.	3.41	1.08	1.00	5.00
		Perform measurements.	4.30	.81	1.00	5.00
		Social				
		Build and manage a team.	4.08	.86	1.00	5.00
		Carry out interdisciplinary projects.	3.86	.97	1.00	5.00
		Organize and lead projects.	4.00	.89	1.00	5.00
		Hold meetings with colleagues from other departments.	4.10	.76	2.00	5.00
		Lead a research group.	4.04	.89	1.00	5.00
		Conventional				
		Plan and manage finances.	3.45	1.04	1.00	5.00
		Do administrative tasks.	3.26	1.08	1.00	5.00
		Artistic				
		Work on inventions.	4.17	.85	2.00	5.00
		Develop measurement methods.	3.97	.82	2.00	5.00
		Construct experimental equipment.	4.17	.79	1.00	5.00
		Teaching				
		Accompany students' theses.	3.73	.90	2.00	5.00
		Supervise students.	3.65	.97	1.00	5.00

Note*:* items with an asterisk (*) were coded reversely.