# Development and validation of the Autonomy Support Situational Judgment Test for Preschool Teachers

**Jianye Fu**[1] | **Fang Jin**[1*]

1 College of Preschool and Elementary Education, Shenyang Normal University, Shenyang, Liaoning, China

**Received:** 24 September 2026 | **Revised:** 26 October 2026 | **Accepted:** 26 October 2026

**ABSTRACT**

Autonomy support is crucial to the development of children's autonomy and intrinsic motivation. Existing evaluation methods face threefold contradictions: convenience versus authenticity, objectivity versus generalizability, and depth of interpretation versus dynamic correlation. To address this, this study developed a set of situational judgment tests for preschool teacher autonomy support. First, we established three dimensions: high control (HC), slight control (SC), and high autonomy (HA). Second, we collected real teaching situations through event sampling observation and selected seven typical situational events using the Delphi method, and 21 behavioral response items were designed. After testing 300 preschool teachers, the results showed that the reliability estimate of the weighted composite score calculated by Mosier's formula was 0.87. Confirmatory factor analysis showed that the model fitted well, and the content validity was verified by expert review and theoretical framework. These findings provide preliminary evidence for the reliability and validity of the tool for assessing preschool teachers' autonomy support. Its utility for teacher training and self-diagnosis requires further intervention and longitudinal research.

**KEYWORDS:** Preschool teachers; Autonomy support; Situational judgment test; Test development

**DOI:** https://doi.org/10.29329/pedper.2026.217 | Vol. 5, No. 2 (2026) | pp. 551-565

## 1. Introduction

Against the backdrop of global educational reforms shifting from traditional standardized instruction to a personalized learning paradigm, the international education community has increasingly emphasized the critical role of teacher autonomy support (TAS) in cultivating students' core competencies and lifelong learning capabilities. This shift reflects the amalgamation of constructivist learning theories and humanistic educational philosophies, advocating that teachers create autonomy-supportive learning environments to stimulate students' intrinsic motivation and critical thinking, and thus promote the return of educational practices to their essential learner-centered orientation. Autonomy support can help teachers overcome the constraints of traditional authoritative teaching, stimulate their teaching innovation vitality, and enhance their professional wellbeing. At the same time, the improvement of teacher autonomy support level also helps to build a more active teacher-student relationship and promote students' autonomous learning and holistic development. This has far-reaching implications for promoting the optimization of the educational ecology and achieving sustainable development of education. Autonomy support is an important

Open access under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/)
**Corresponding author:** jinfang1997@163.com

concept in self-determination theory. It refers to the support that meets the autonomy of individuals, allowing individuals to act in a way that conforms to their own values, goals and interests, thereby enhancing their intrinsic motivation and self-regulation ability (Deci & Ryan, 2002). Autonomy support strategies can meet students' individual learning needs and promote their learning engagement (Wang, 2026). The early childhood stage is a critical period for the growth and development of individuals, during which the autonomy support of preschool teachers is particularly important. Autonomy support of preschool teachers can be understood as interactive strategies that support the development of children's intrinsic motivation and autonomy. It is manifested in behaviors such as providing opportunities for choice, using noncontrolling language, acknowledging children's emotional needs, and encouraging self-decision making. The OECD's Starting Strong V: Transitions from Early Childhood Education and Care to Primary Education makes a compelling case for preschool teachers to create interaction environments that allow children full autonomy to express and explore (Organisation for Economic Co-operation and Development [OECD], 2017). This is highly consistent with the concept of autonomy support, highlighting the important position of autonomy support by preschool teachers in early childhood education. Such supportive educational practices can not only promote the development of autonomy and intrinsic motivation in young children, but also provide them with a favorable educational environment for their overall development. Therefore, preschool teachers should actively adopt autonomy support strategies when implementing educational activities in order to promote the healthy growth and all-round development of children.

In educational practice, researchers adopt a variety of ways to evaluate teacher autonomy support behavior. Questionnaire survey is simple to operate and suitable for large-scale collection and statistical analysis, but it is prone to the influence of subjects' subjective biases and social expectation effect (Cui & Bao, 2025). Classroom observation method provides authentic data to improve teaching by directly recording the details of teaching interactions, but it has limitations such as high cost, long duration, and interference with the behavior of teachers and students (Villegas et al., 2025). Event sampling observation has the advantages of authenticity, systematicness and strong pertinence, but it is limited by the subjective judgment of the observer and the lack coverage (Zachariou & Bonneville-Roussy, 2024). In order to make up for the shortcomings of a single method, a mixed method such as explanatory sequential design has emerged. It combines quantitative and qualitative evaluation to ensure the reliability and validity of the data, and also allows for an in-depth exploration of the complex mechanisms of influence. However, the two types of data are prone to disconnection during the integration process, making it difficult to systematically present the dynamic mechanisms of teacher autonomy support behaviors in real scenarios. (Wang et al., 2024). From this perspective, the current evaluation system reveals three types of contradictions: the contradiction between convenience and authenticity, the contradiction between objectivity and generalizability, and the contradiction between the depth of explanation and dynamic correlation. Therefore, it is imperative to develop an evaluation tool that can fully and accurately reflect teacher autonomy support behaviors, so as to break through the triple contradiction of the current evaluation system.

Situational judgment test (SJT) provides a new idea to solve the above three-fold contradiction. SJT is an effective tool to assess individuals' judgment and decision-making ability in specific situations, through presenting problems in concrete situations and asking the testees to choose the best solution (Firouzabadi & Keshmiri, 2025). In the evaluation of kindergarten teacher autonomy support behaviors, situations can be designed that are close to real teaching scenarios. For example, when children encounter difficulties during autonomous play or do not follow the rules during collective

activities, allow teachers to choose the most appropriate way to deal with them. This method can simulate real scenarios, reduce subjective biases and expectation effect, and improve the authenticity of data, breaking through the contradiction between convenience and authenticity. It can be implemented on a large scale, reducing costs and time investment, and improve the objectivity of evaluation through standardized situations, breaking through the contradiction between objectivity and generalizability. Through the analysis of teachers' responses in different situations, it can explore the performance and motivation of autonomy support behaviors, show the changes of behaviors in dynamic situations, and break through the contradiction between explanatory depth and dynamic correlation.

At present, situational judgment tests have gradually been applied to the field of preschool education. Nilsen et al. (2020) used situational judgment questions (SJQs) as a core tool to assess the quality of the process of preschool education in the field trial of the OECD's 2018 "Teaching and Learning International Survey of Strong Starts", covering three dimensions: encouraging prosocial behavior, behavior management, and supporting child-led play. Li et al. (2022) developed a situational judgment test of analysis and response skills for preschool teachers (SJTs-PTSAR), which includes 16 scientific education-related situations to examine the skills of teachers in analyzing and responding to the developmental trajectory of children's pre-concepts. Alsaeed and Mohammad (2023) used a situational judgment test with 15 authentic mathematics teaching situations as a core assessment tool to carry out a 15-week situational professional development program, which effectively enhanced the student-centered teaching practice ability of preschool normal students. Thohiroh et al. (2025) developed a Mindful Parenting in Preschoolers Situational Judgment Test (MPP), which retained 30 items and three dimensions of awareness, sympathy, and non-reactivity after testing, and was suitable for parents of preschool children aged 3–5, and was significantly negatively correlated with parenting stress. Murano (2019) developed a comprehensive assessment tool including a situational judgment test based on the Big Five personality theory framework in her doctoral thesis, which provided new ideas for the assessment of social and emotional skills in preschool children. This indicates that the situational judgment test plays an increasingly important role in the field of preschool education, but the situational judgment test for the autonomy support of preschool teachers has not yet been developed.

Therefore, based on the existing mature scale of autonomy support, this study follows the scientific development requirements of situational judgment tests, and compiles a set of situational judgment tests that can be used for scientific evaluation of the autonomy support level for preschool teachers.

## 2. The Process of Test Construction

### 2.1. Determine the Measurement Structure

To identify the measurement structure, it is first necessary to clarify the structural dimensions of autonomy support. Deci et al. (1981) proposed the important concept of autonomy support and developed The Problems in Schools Questionnaire (PIS) to examine teacher autonomy support. The questionnaire uses eight teaching situations suitable for primary and secondary schools, lists four categories of incentive behaviors that might occur in each situation, namely high autonomy, moderate autonomy, moderate control, and high control, and the teachers rate their tendency to adopt each teaching behavior (Halvari et al., 2022). Reeve et al. (1999) found among the four types of teacher incentive behaviors proposed by Deci, the score of moderate autonomy was positively correlated with moderate control and high control, and moderate autonomy tended to be more controlling. Therefore, they adopted the method of deleting the response items related to moderate

autonomy to modify the questionnaire, and conducted validity testing on the modified PIS questionnaire, and showed that it had a high validity. Later, Chinese scholar Wang (2018) found that among the four types of teacher incentive behaviors proposed by Deci, distinction between moderate autonomy and moderate control was small. Theoretically, the key difference between autonomy support and control in Self-Determination Theory lies not in the degree of teacher involvement, but in whether children's autonomy needs are met and whether they are regarded as decision-making agents. Although moderate autonomy and moderate control differ in control intensity, both essentially center on teacher-preset goals, instructional frameworks, and problem-solving paths; teachers merely guide children in a gentler, indirect, or heuristic manner, without truly returning the right to choose and solve problems to the children. In contrast, high autonomy emphasizes respecting children's autonomy in both emotion and behavior, providing opportunities for choice, and encouraging children to think from multiple perspectives and attempt to solve problems themselves. Therefore, merging moderate autonomy and moderate control into "slight control" is not a simple reduction of dimensions, but a restructuring of the control-autonomy continuum based on conceptual homogeneity and empirical discriminability: high control represents strong control, mild control represents weak control but still belongs to the control orientation, and high autonomy represents true autonomy support. Compared to the original four-category model, the three-dimensional model (HC-SC-HA) is more appropriate: first, its conceptual boundaries are clearer, avoiding semantic overlap between "moderate autonomy" and "moderate control" or "high autonomy"; second, it has better empirical discriminability, reducing the misjudgment of gentle control as autonomy support in teacher self-reports due to social desirability; third, it better fits the measurement requirements of situational judgment tests, with three response items—high control, slight control, and high autonomy—set for each situation, facilitating clear behavioral orientation judgments and weighted scoring. Therefore, this study took the questionnaire compiled by Deci as the basis, referred to the modification method of Reeve et al., combined the two items of moderate control and moderate autonomy into slight control, and determined that high control, slight control and high autonomy were measurement structure (see Table 1 for details), and the situational judgment test that can be used to evaluate the control or autonomy orientation adopted by preschool teachers in interaction with children was compiled.

**Table 1** ***Measurement structure of the autonomy support questionnaire for preschool teachers***

| Structure | Connotation |
|---|---|
| **High control (HC)** | Teachers take charge or solve problems in a punitive/rewarding way. *E.g.,* giving the answers to the questions directly or demanding requirements in commanding way. |
| **Slight control (SC)** | Compared with high control, the lower degree of control can provide necessary help for young children, but essentially still solve problems according to the teachers' way of thinking. *E.g.,* inspire children through communication, encourage children to use references to create. |
| **High autonomy (HA)** | Teachers respect the autonomy of young children in emotion and behavior, encourage them to consider multiple aspects of problems, give them opportunities to choose, and try to find solutions by themselves. *E.g.,* listen to the ideas of young children and encourage them to try boldly. |

## 2.2. Collection and Screening of Situational Events

### *2.2.1. Collection of Situational Events*

Situational events provide the testees with the background and basis for assessment in the situational judgment test, stimulate their cognition and response, and serve as the standard for evaluation and comparison. In the situational judgment test of autonomy support for preschool teachers, using segments of kindergarten life that teachers are familiar with as measurement questions is more likely to elicit the true reactions of the subjects and measure the true results. To this end, the researchers went deep into the kindergarten and collected situational events using the event sampling observation method. First, the observation situations were determined. Based on the study of preschool education books, journals and kindergarten cases, seven situations in which teacher autonomy support often occurs were sorted out: indoor centralized teaching activities, outdoor centralized teaching activities, game activities, learning activities, post-activity sorting, meals (snacks and full meals), and other activities. Subsequently, a month-long follow-up observation was carried out in the small, middle and large classes of C kindergarten in S city. Two hours each morning and afternoon were selected for observation, and the situations in which teachers interacted with young children were recorded, including the causes, processes, results, and the reactions of teachers and young children in this event. Finally, the collected situational events were collated, the situational events that did not conform to the content of teacher autonomy support for young children or were meaningless were deleted, and the typical ones that could better reflect the autonomy support were selected for linguistic processing and integration, and finally a total of 24 situational events were collected.

### *2.2.2. Screening of Situational Events*

Typicality and commonality are the criteria for screening situational events in the situational judgment test, which also affect the validity and reliability of the test (Chen & Burns, 2025). Therefore, this study selected five graduate students in preschool education with more than half a year of internship experience and 15 preschool teachers with more than three years of teaching experience to form an evaluation group, asking them to evaluate the typicality and commonality of the situational events collected, which were divided into two links. The first link was to judge the typicality of the situational events, and the evaluation group discussed whether the situational events collected were representative, deleting the situational events that the evaluation group couldn't reach a consensus and the low representative ones. The second link was a 7-point scoring evaluation (1= "least common" and 7 = "most common") on whether the situational events had commonality by the group members. The larger the number, the more common the situational events are. After the evaluation by the evaluation group, evaluation results were statistically analyzed, and the situational events with a total score below the average and a high standard deviation in the same situation were eliminated.

Finally, after two rounds of screening, nine scenarios with good typicality and commonality were obtained (The standard deviation range is 1.099–1.525), and the scenario language description was modified and integrated to the official problem scenario stem content. The scenario problem content (excerpt) is shown below (Table 2).

**Table 2** ***Summary of situational events (excerpt)***

| Classification | Problem Situation |
|---|---|
| **Teaching activity** | In an art activity, the teacher asked children to draw a slide they had seen with their favorite colors. Everyone began to focus on drawing. The teacher noticed that Du Du was sitting still and had not started. The teacher asked, and Du Du said, "Teacher, I don't know how to draw a slide." |
| **Game activity** | Longlong was carefully building his sky garden in the construction area and had almost finished. Xiao Yu rushed over and accidentally knocked it down. Longlong became very angry. |
| **Learning activity** | The teacher bought a crab from the market. Zibo loved it and went to see it first thing every day. Today Zibo found that the crab had shed its shell and asked many questions about it, refusing to leave even after class started. |
| **Life activity** | Mingming, a child in the junior class, disliked vegetables and ate very little. At lunch, there were broccoli and spinach. As soon as Mingming sat down, he pushed the vegetables to the edge of the plate with his spoon and frowned, saying, "I don't want to eat this!" |

## 2.3. Collection and Screening of Behavioral Response Items

### *2.3.1. Collection of Behavioral Response Items*

Behavioral response items are multiple selectable behavioral response options listed for each specific work or life situation description in the situational judgment test, representing the different actions or coping methods that individuals might take in that situation. To obtain discriminative and authentic behavioral responses in nine situational events, a semi-structured interview method was used to collect behavioral response items. The identified nine situational events were designed into the interview content, and 15 preschool teachers were interviewed, asking the teachers to answer the practices they might take in these situations, rather than what they thought was the best practice. After the interview, the possible coping measures of preschool teachers in this situation were recorded, and then the interview data organized into response options according to the three types of behaviors: "high control", "slight control", and "high autonomy". After preliminary organization, five behavioral response items were obtained for each situational event, totaling 45 behavioral response items for nine situational events.

### *2.3.2. Arrangement and Screening of Behavioral Response Items*

#### *2.3.2.1. Arrangement of Reaction Options*

The researchers combined the behavioral responses collected, removed the behaviors that had nothing to do with the situation, and then classified the items of the behavioral responses that had preliminarily collated according to the three dimensions of "high control", "slight control", and "high autonomy". For example, when asked "what would you do if a child can't draw a slide in an art activity", five teachers gave their responses based on their usual practice (Table 3).

**Table 3** ***Examples of teacher responses***

| Teacher | Response |
|---|---|
| T1 | I would not reproach him for not doing as others do, but I would first ask him where he is at fault, and then encourage him: What do you like to draw? Go ahead and try. |
| T2 | Before drawing the slide, I will show the children a lot of pictures of slides, then let the children talk about what parts the slide is made of, then let them draw. When drawing, one or two demonstration pictures will be put up. If there are still children who don't know how to draw, I will let them watch the slide in the demonstration drawing, let them talk about what the slide looks like, and then let them draw. |
| T3 | I will guide him: Look, what does the slide in our kindergarten look like? Some are like rectangles, some are like circles, they are built together, with many beautiful colors, so let's draw a rectangular slide first! I can hold his hand and lead him to draw. |
| T4 | I will use language cues, such as: What does the slide in our kindergarten look like? What color is it? Try to draw it. |
| T5 | I would ask him: why can other children do it and you can't? Do you still want to draw a slide? If you still want to draw, then draw according to the slide in the yard. |

It can be seen from the table that when T1 found that children could not draw a slide, she did not blame them, but guided them to express their needs in a questioning way and encouraged them to try boldly. Emotionally and behaviorally, she understood and respected children's autonomy and should be attributed to the high autonomy dimension. When T5 found that children could not draw a slide, she not only blamed them but also required them to draw according to the slide in the kindergarten in a warning tone. In the process of solving the problem, the teacher always occupied the dominant position, neglected the children's will, and her blame, warning and demands put pressure on the children. Finally, she forced the children to complete the painting according to the way she stipulated, so she was attributed to the high control dimension. T2, T3 and T4 inspired the children by talking about the appearance of the slide. Although they gave the children a chance to discuss and provided necessary help for the children, they did not explore the reasons why the children could not draw the slide. Essentially, they still used the teacher's thinking to solve the problem. Especially the behavior of T2 and T3 had obvious control. They can be attributed to the slight control dimension.

#### *2.3.2.2. Screening of Reaction Options*

The Delphi method was used to further screen the items of response, and an expert group was formed by inviting three preschool education experts and two preschool teachers, each with more than three years of teaching experience, for evaluation. First, the members of the expert group were trained to know the meaning and characteristics of high control, slight control and high autonomy, in order to the way for the subsequent evaluation. Then, the experts were divided into two groups, and the two groups of experts evaluated whether the items of response could reflect the autonomy support and control of preschool teachers and whether they could effectively distinguish between high control, slight control and high autonomy. After the first round of evaluation, the options agreed upon by the two groups of experts are retained, and the options deemed by the experts to lack discriminative power are excluded. Options with conflicting opinions proceed to the second round of discussion, where the expert panel will decide whether to delete or modify them after further deliberation. Finally, combined with the results of the expert evaluation, three behavioral response items were formed in each problem situation. For example, in the situation where children do not know how to draw a slide, the response items are: let children refer to the example picture to draw, and tell children that they

can be rewarded with a small sticker for drawing a slide (high control); talk to children about what shape the slide in the kindergarten is, let children try to draw what they think (slight control); ask children why they can't draw, and encourage children to create boldly (high autonomy). By this stage, the situational judgment test with nine situational events, three response options for each event, and a total of 27 behavioral response items was obtained.

### 2.4. Content Review and Scoring Method

#### *2.4.1. Review of Test Content*

After the above steps, the first draft of the questionnaire on autonomy support for preschool teachers was determined. After the first draft was formed, three preschool education experts and two excellent preschool teachers were invited to give suggestions on whether the test itself was in line with the theoretical conception, the correlation between situational events and autonomy support for preschool teachers, and the discrimination between items of high control, slight control and high autonomy. At the same time, in order to test whether the text expression of the questionnaire was reasonable, five graduate students majoring in preschool education were invited to judge whether there was ambiguity in the text expression of the measurement and to give suggestions. Finally, the test content was perfected according to their opinions.

#### *2.4.2. Determination of Scoring Method*

The scoring of test items is an important step in quantifying the test-takers' responses, focusing on the design of appropriate scales and scoring methods to measure the test-takers' responses to the items (Downing, 2006). The questionnaire requires teachers to read each question situation carefully, then consider three possible answers' alignment with their own practices in their daily work, and use a 7-point Likert scale to score each item to indicate the extent to which they believe the answer aligns with their actual practice, with 1 indicating "not consistent", 4 indicating "possibly consistent", and 7 indicating "very consistent". The scale consists of 27 items, of which nine items are highly controlling, belonging to the high control scale (HC); nine items are lightly controlling, belonging to the slight control scale (SC); and nine items are highly autonomous, belonging to the high autonomy scale (HA). The research refers to the results of Deci et al., and the questionnaire scores are weighted to ensure that the questionnaire results can effectively reflect the level of autonomy support of preschool teachers. First, the scores of the nine items in each subscale are averaged to obtain the score of each subscale, ranging from 1–7 points. Then, the scores of the high control scale (HC), the slight control scale (SC), and the high autonomy scale (HA) are weighted as -2, -1, and 2, respectively, and then summed to obtain the total scale score. The total scale score is: 2HA-SC-2HC, the higher the score, the more inclined the preschool teachers are to autonomy support (Deci et al., 1981). The rationale for this weighting formula is as follows: First, high autonomy and high control represent the two ends of a continuum, and are therefore assigned weights of 2 and −2 respectively to reflect their opposing contributions to autonomy support. Second, moderate autonomy and moderate control lack sufficient differentiation, and moderate autonomy is closer to a control orientation; therefore, this study merges the two into slight control and assigns it a weight of −1, representing "weak control" rather than autonomy support. Third, using a weighted sum instead of a simple summation preserves the theoretically different directions and intensities of the three dimensions, avoiding treating controlling behaviors and autonomy-supportive behaviors as equivalent.

## 3. Test Administration

### 3.1. Pretest Administration

#### *3.1.1. Subject Selection*

The pre-questionnaire was randomly distributed to preschool teachers within the scope of Liaoning Province, and 118 questionnaires were collected. After excluding invalid and incomplete information, a total of 114 valid questionnaires were obtained, with an effective rate of 96.6%.

#### *3.1.2. Project Analysis*

The project analysis was conducted using the decisive extreme value method, where the top 27% and bottom 27% of the total scores were set as the high and low groups. Then, the independent sample t-test was applied to compare the differences between the high and low groups in each item. In the project analysis, items with a *CR* value ≥ 3 and a significance level of .01 were considered to be items with discriminatory power. The test results showed significant differences existed between the high and low groups in all items, but one item had a *CR* value of -2.63, with an absolute value less than 3, and it was not discriminative, so it was deleted. Calculate the point-biserial correlation between each item and the total scale score after deducting that item; any item with a correlation coefficient that is significant or is below 0.4 is considered a lowly correlated item. Under this criterion, no new items were deleted. In this test, the highly controlled and lightly controlled items are assigned negative weights during scoring (HC=-2, SC=-1) and are considered control items. During item analysis, control items are negatively or weakly correlated with the total score, which is the theoretically expected result. Therefore, in the item-total correlation analysis, control items are not simply deleted due to their negative correlation with the total score, but are comprehensively evaluated in conjunction with context-based discrimination tests. The questionnaire was a situational judgment test, with each question situation containing three behavioral response items, representing high control, slight control, and high autonomy, respectively. If the correlation between the control tendency and autonomy tendency items in a situation reached a significant level, it indicated that the situation and behavioral response items cannot distinguish the test-takers' control or autonomy support tendency. Therefore, two situations that did not meet the theoretical assumptions were deleted, and seven problem situations with discriminatory power were retained. The formal questionnaire included seven problem situations and 21 behavioral response items, covering the three dimensions of high control, slight control and high autonomy.

Wu (2010) proposed that the project compiled based on the theoretical basis can omit the exploratory factor analysis. In this study, the questionnaire of situational judgment test clearly defined the measurement structure dimensions in advance, and the questions were constructed around specific situational tasks and evaluation dimensions. The response options and items of the questionnaire were reviewed by the Delphi method, and the corresponding relationship between the questions of the questionnaire and the content to be measured was clear, so the exploratory factor analysis was omitted.

### 3.2. Formal Administration

#### *3.2.1. Subject Selection*

According to the principle of convenient sampling, preschool teachers in Shenyang, Fuxin, Chaoyang, Dalian in Liaoning Province and Xi'an in Shaanxi Province and Beijing were selected as the research objects. A total of 308 questionnaires were distributed, and 303 were actually collected. After excluding invalid questionnaires, 300 valid questionnaires were obtained, with an effective recovery rate of 97.4%. Regarding the determination of the sample size, this study referred to the general

requirements for scale development and confirmatory factor analysis. First, the situational judgment test contains 21 behavioral response items, and based on a 1:10 ratio of items to sample size, the required sample size is approximately 210. Second, confirmatory factor analysis usually requires a sample size of over 200. A sample of 300 can meet the stability requirements for model estimation. In terms of gender, there are 7 male teachers (2.3%) and 293 female teachers (97.7%). In terms of age, there are 8 teachers aged 20 and below (2.7%), 192 teachers aged 21–30 (64.0%), 75 teachers aged 31–40 (25.0%), 22 teachers aged 41–50 (7.3%), and 3 teachers aged 51 and above (1.0%), with the sample mainly consisting of young and middle-aged teachers. In terms of teaching experience, there are 39 teachers with less than 1 year (13.0%), 149 teachers with 1–5 years (49.7%), 65 teachers with 6–10 years (21.7%), 27 teachers with 11–15 years (9.0%), and 20 teachers with over 16 years (6.7%), covering both novice and experienced teachers.

### *3.2.2. Reliability Analysis*

The internal consistency reliability of each dimension of the autonomy support situational judgment tests for preschool teachers and the total questionnaire was examined by using Cronbach's $\alpha$ coefficient. The results showed that the $\alpha$ coefficients of each dimension of the questionnaire ranged from 0.74 to 0.89. The formal scoring adopts the weighted composite score of 2HA−SC−2HC, whose reliability is estimated based on Mosier's (1943) weighted composite reliability formula, combined with the reliability of the three subscales and the inter-factor correlations obtained from confirmatory factor analysis, resulting in 0.87. This indicates that the weighted composite score, which matches the actual scoring procedure, has good reliability and can stably reflect the level of autonomy support among preschool teachers. The specific results are shown in Table 4.

**Table 4** ***Reliability estimates of the formal questionnaire***

| | HC | SC | HA | Total Test |
|---|---|---|---|---|
| **Reliability estimate** | 0.89 | 0.74 | 0.80 | 0.87 |

*Note.* Cronbach's α was calculated for each subdimension (HC, SC, HA). The reliability of the weighted composite score (Total Test) was estimated using Mosier's formula.

### *3.2.3. Construct Validity*

Confirmatory factor analysis was conducted using AMOS 23.0 software to detect whether the questionnaire structure conformed to the theoretical conception. Considering that each question in this test contains three behavioral response items representing high control, slight control, and high autonomy respectively, the three items within the same situation may share situational context, thereby generating situation-related shared variance. Therefore, the study allows for partial residual correlation within the same context to control the impact of context-shared variance on model fitting. With high control, slight control, and high autonomy as latent variables, each latent variable included seven observed variables respectively, which was used as the basis for testing the model of autonomy support for preschool teachers, and the results are shown in Table 5. When the model fit is judged by using confirmatory factor analysis, it mainly refers to the values of $\chi^2/df$, RMSEA, GFI, CFI, and so on. The closer the value of $\chi^2/df$ is to 1, the higher model fit to the data, and the value of $\chi^2/df$ is less than 2, indicating that the model fit is excellent, between 2–5 is acceptable; the value of RMSEA below 0.05 indicates that the model fit is excellent; the GFI, CFI, IFI change range is between 0–1, and the closer the values of these three are to 1, the better the model is fitted. It can be seen from Table 5 that all the indexes of the model meet the statistical requirements. In which, $\chi^2/df < 2.00$, RMSEA < 0.05, value of GFI is close to 0.85, and the values of CFI and IFI are greater than 0.90. This shows that the questionnaire is reasonable and the model fit is good.

**Table 5** *Main model fit indices in the confirmatory factor analysis of the formal questionnaire*

| Indicator | *df* | $\chi^2$ | $\chi^2/df$ | RMSEA | GFI | CFI | IFI |
|---|---|---|---|---|---|---|---|
| **Numeric** | 184 | 261.72 | 1.42 | 0.04 | 0.83 | 0.91 | 0.91 |

#### *3.2.4. Content Validity*

This study ensures the questionnaire has good content validity through a systematic development process. First, the collection of situational events is based on real kindergarten observations, covering typical interaction scenarios in daily teaching, and is ensured to be highly relevant to kindergarten teacher autonomy support behaviors by a panel of judges' selecting for typicality and commonality. Behavioral response items were collected by semi-structured interviews to obtain real teacher feedback, and the options were discriminated by an expert panel using the Delphi to ensure that the response items could effectively distinguish between the three types of behaviors: high control, slight control, and high autonomy. In addition, the draft of the test was reviewed by preschool education experts, front-line teachers, and graduate teams from multiple dimensions such as theoretical fit, situational relevance, option discrimination, and clarity of expression, and it was repeatedly revised based on the feedback. The final form of the situational judgment test not only fits the actual teaching situation but also accurately reflects the theoretical framework of autonomy support, fully guaranteeing the representativeness and adequacy of the test content.

It is particularly worth noting that this test adopts a teacher self-report method, measuring teachers' self-reported behavioral tendencies rather than the autonomy-supportive behaviors actually observed in the classroom. Discrepancies may exist between self-reported behavioral tendencies and actual classroom behaviors. Future research should combine multiple methods such as classroom observations, student perceptions, or external expert ratings to further examine the criterion-related validity and ecological validity of the test.

## 4. Discussion

### 4.1. Structure of the Test

Autonomy support situational judgment test for preschool teachers is composed of seven short situational events that describe typical problems that may occur in a kindergarten. Three possible responses to each problem situation, representing the three tendencies of high control, slight control, and high autonomy, follow the problem situation, respectively. The order of the three responses in each problem situation is balanced, with no precedence given to any one.

A large amount of research on the structure of teacher autonomy support has been conducted by scholars, providing theoretical basis for the dimensional division of test. In previous studies, there are two paths to the dimensional division of autonomy support behaviors: one is overall evaluation based on a single dimension, which regards teacher autonomy support as a continuous variable and evaluates its intensity through a Likert scale (Monacis et al., 2023); the other is to construct a multi-dimensional structure from the perspective of support content (Zhao et al., 2025). However, both of these two classification methods have limitations in the development of situational judgment tests. Compared with them, the three-dimensional model of high control, slight control, and high autonomy adopted in this study is more practical. The initial structure of the test is based on "The Problems in Schools Questionnaire" (PIS) by Deci et al., which divides teacher autonomy support into four categories: high control, moderate control, moderate autonomy, and high autonomy. However, subsequent studies have found that there is conceptual confusion between moderate autonomy and

moderate control, resulting in insufficient discriminant validity. Thus, this study referenced the practice of Reeve et al. and deleted the moderate autonomy dimension, and finally formed a three-dimensional structure of high control, slight control, and high autonomy. The situational events cover seven typical scenarios such as indoor and outdoor activities, games, and life links, which fully map the spectrum of behaviors from control to autonomy of preschool teachers. Meanwhile, the content validity is further reinforced through systematic screening of expert evaluations by the Delphi method and behavioral response items. Therefore, the test has a good theoretical structure and educational practical adequacy.

### 4.2. Validity of the Test

The questionnaire had an internal consistency coefficient of 0.87, and the $\alpha$ coefficients of each dimension ranged from 0.74 to 0.89, reaching an acceptable level, indicating that the test had good reliability. From the results of the confirmatory factor analysis, the three dimensions of high control, slight control and high autonomy fit the theoretical conception and have good structural validity. The research distributed questionnaires to 300 teachers and found that preschool teachers had the highest score in the dimension of high autonomy, and there was not much difference among teachers; teachers had the lowest score in the dimension of high control, and there was a large difference among teachers. These are consistent with the conclusions of previous studies, which shows that the assessment tool constructed in this study can effectively measure the level of autonomy support of preschool teachers (Wang, 2018). These results preliminarily support that the assessment tool constructed in this study can be used to measure the level of autonomy support among preschool teachers, but its diagnostic utility, cross-sample stability, and predictive validity for teachers' actual behavior still require further research verification.

### 4.3. Potential Value of the Test for Teacher Training

Situational judgment tests are based on the assumption of behavioral consistency, which refers to the consistency of individuals' past behaviors being a predictor of future behaviors across situations. In other words, the behavioral options provided by situational judgment tests serve as samples of overall behavioral performance, and the individuals' judgment and selection of the samples provided by the test is a prediction of their current level of behavior and future tendencies. Therefore, situational judgment tests are professional judgments of teachers' concepts and behaviors in the process of implementing nursery education, and it is feasible to use them to examine the possibility of effective support provided by teachers in their interaction with young children. At the same time, the establishment of this assessment tool provides the possibility for teachers' self-diagnosis and improvement. As Landy (2007) said, situational judgment tests are "a kind of alchemy" that can reduce negative impacts and improve effectiveness. The situational judgment tests developed by this study are based on the practice of nursery education in kindergartens and are evaluated by teachers themselves, which can promote teachers' self-evaluation and reflection combined with their own practice experience.

Han et al. (2011) pointed out that teacher autonomy support behavior can be trained. However, previous training for preschool teachers often focused on teachers' teaching ability, and the development of children's autonomy seemed to be emphasized as a concept attached to various training courses or activities, which led to teachers still not knowing how to provide effective support when facing actual problems. Therefore, the application of autonomy support situational judgment tests questions for preschool teachers as training materials in teacher training, and the clarification of abstract concepts into action frameworks that can be referred to, helps to promote the

transformation of training effectiveness and thus improve the level of autonomy support for preschool teachers.

## 5. Reflection and Prospect

This study constructed the "Autonomy Support Situational Judgment Test for Preschool Teachers" based on self-determination theory, which has produced preliminary evidence in test development and reliability and validity examination, and suggests potential practical application. However, there are still the following limitations: First, the sample is mainly concentrated in some cities, and the geographical coverage is limited. Although convenience sampling is feasible in data collection, it may limit the generalizability of the research findings to preschool teachers and kindergartens in other regions. In the future, the sample range needs to be expanded to enhance the applicability of the tool. Second, although the test situation has been strictly selected, it only covers seven typical situations. In the future, it can be combined with the new educational models of kindergartens (such as STEM activities, mixed-age education) to supplement various situations and improve the dynamic adaptability of the test. Third, the classification of behavioral response items depends on the subjective judgment of experts. In the future, the cognitive mechanisms of teacher decision making can be verified by eye tracking or neurophysiological indicators to strengthen the objectivity of classification.

Looking ahead, the following three aspects can be further investigated: First, the correlation between teacher autonomy support level and children's developmental indicators (eg. creativity, social adaptability) can be longitudinally tracked, revealing its long-term educational effect; Second, teacher training courses based on test results can be developed combined with situational simulation and reflective log, to explore the integrated "assessment-feedback-intervention" model; Third, cross-cultural comparison of the differences in autonomy support behaviors of preschool teachers in different regions can be made to distill universal and culturally specific strategies, providing empirical basis for educational policy formulation. In addition, the integration of artificial intelligence technology can be attempted to construct a dynamic situational generation system, exploring the possibilities of personalized feedback and real-time assessment.

## 6. Declarations

### 6.1. Conflict of interest

The authors declare no competing interests.

### 6.2. Funding

This study was supported by the General Project of the National Social Science Fund of China (Education) in 2024, Project No. BQA240217.

### 6.3. Author contributions (CRediT)

Jianye Fu: Conceptualization; Methodology; Writing—original draft. Fang Jin: Investigation; Data curation; Writing—review & editing. All authors have read and approved the final version of the manuscript.

### 6.4. Data availability statement

The data supporting the findings of this study are not yet publicly available.

### 6.5. Ethics approval

This study was approved by the Shenyang Normal University Research Ethics Committee (approval date: 30 January 2026; reference number: 20260130) and was conducted in accordance with its ethical guidelines. Before the interviews and the completion of the questionnaires, written informed consent was obtained from all participating teachers, who were informed of the purpose of the study, the anonymity of their responses, and the voluntary nature of their participation. Classroom observations were conducted after consent had been obtained from the kindergarten

principal and the class teachers, and written informed consent was obtained from the parents or legal guardians of the children observed. Normal teaching activities were not disturbed, and the observation records were used only for research analysis and did not involve any evaluation of the teachers or children.

### 6.6. Use of artificial intelligence (AI) tools

No generative artificial intelligence (AI) or AI-assisted tools were used in the design of the study, data collection, data analysis, interpretation of the findings, or the writing and editing of this manuscript. The authors take full responsibility for the content of the publication.

### 6.7. Acknowledgements

The authors would like to thank all the teachers who participated in this study, as well as the kindergarten directors who facilitated the recruitment of participants in the survey.

## References

Alsaeed, M. S., & Mohammad, A. D. A. (2023). Early childhood student teachers engaging in a scenario-based professional program: The case of early mathematics. *Cogent Education*, *10*(2), 2281747.

Chen, X., & Burns, G. N. (2025). Different paths, same destination? Comparison of two approaches to developing situational judgment tests for cross-cultural competence. *Applied Psychology*, *74*(4), e70024. https://doi.org/10.1111/apps.70024

Cui, S., & Bao, Q. (2025). The relationship between family violence typology and university student cyberbullying: The mediating role of deviant peer affiliation and the moderating role of homeroom teacher autonomy support. *Acta Psychologica*, *261*, 105962. https://doi.org/10.1016/j.actpsy.2025.105962

Deci, E. L., & Ryan, R. M. (Eds.). (2002). *Handbook of self-determination research*. University of Rochester Press.

Deci, E. L., Schwartz, A. J., Sheinman, L., & Ryan, R. M. (1981). An instrument to assess adults' orientations toward control versus autonomy with children: Reflections on intrinsic motivation and perceived competence. *Journal of Educational Psychology*, *73*(5), 642–650. https://doi.org/10.1037/0022-0663.73.5.642

Downing, S. M. (2006). Twelve steps for effective test development. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 3–25). Lawrence Erlbaum Associates.

Firouzabadi, N. F., & Keshmiri, F. (2025). Situational judgment test in pharmacy education: Assessing professionalism capability among students. *BMC Research Notes*, *18*(1), 128. https://doi.org/10.1186/s13104-025-07183-6

Halvari, A. E. M., Ivarsson, A., Halvari, H., Dahl, K. E., Olafsen, A. H., Solstad, B. E., ... Williams, G. (2022). Dental hygienists' biopsychosocial beliefs and giving autonomy support in treatment of patients: A self-determination theory perspective. *International Journal of Dental Hygiene*, *20*(2), 193–202. https://doi.org/10.1111/idh.12584

Han, J. T., Ge, M. G., & Zou, C. H. (2011). Teacher autonomy support: Western research and its enlightenment. *Foreign Education Research*, *38*(6), 49–54. (In Chinese)

Landy, F. J. (2007). The validation of personnel decisions in the twenty-first century: Back to the future. In S. M. McPhail (Ed.), *Alternative validation strategies: Developing new and leveraging existing validity evidence* (pp. 409–426). Wiley.

Li, L., Lv, X., Xi, Y., & Guo, L. (2022). Can preschool teachers' accurate analysis of the development trajectories of children's preconceptions ensure their effective response? Evidence from situational judgement tests. *Sustainability*, *14*(18), 11725. https://doi.org/10.3390/su141811725

Monacis, D., Sulla, F., Peconio, G., & Limone, P. (2023). Measuring autonomy support in special needs teachers from a self-determination theory perspective: Validation of the Italian version of the learning climate questionnaire. *Frontiers in Psychology*, *14*, 1183205. https://doi.org/10.3389/fpsyg.2023.1183205

Mosier, C. I. (1943). On the reliability of a weighted composite. *Psychometrika*, *8*(3), 161–168. https://doi.org/10.1007/BF02288700

Murano, D. M. (2019). *Improving measurement and expanding meta-analytic knowledge: Social and emotional learning in elementary and early childhood* [Doctoral dissertation, City University of New York]. ProQuest Dissertations Publishing.

Nilsen, T., Slot, P., Cigler, H., & Chen, M. (2020). Measuring process quality in early childhood education and care through situational judgement questions (OECD Education Working Papers No. 217). OECD Publishing. https://doi.org/10.1787/852602c5-en

OECD. (2017). *Starting strong V: Transitions from early childhood education and care to primary education*. OECD Publishing. https://doi.org/10.1787/9789264276253-en

Reeve, J., Bolt, E., & Cai, Y. (1999). Autonomy-supportive teachers: How they teach and motivate students. *Journal of Educational Psychology*, *91*(3), 537–548. https://doi.org/10.1037/0022-0663.91.3.537

Thohiroh, A. A., Dewi, M. A., Maulina, D., Maulani, N. M., Faradila, R., & Latifah, T. (2025). The development of mindful parenting in preschoolers (MPP) scale. *Jurnal Ilmu Keluarga Dan Konsumen*, *18*(1), 66–79. https://doi.org/10.24156/jikk.2025.18.1.66

Villegas, F. E., van der Steen, S., van Dijk, M., Villegas, D. A. E., & Minnaert, A. (2025). Teacher-student interactions of autistic adolescents: Relationships between teacher autonomy support, structure, involvement and student engagement. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-025-06723-0

Wang, C. M. (2018). Investigation and research on the types and current situation of kindergarten teachers' autonomy-supportive behaviors [Master's thesis, Shenyang Normal University]. (In Chinese)

Wang, M. M. (2026). A study on the relationship between teacher autonomy support, academic buoyancy and high school students' English learning engagement. *New Explorations in Education and Teaching*, *4*(1), 96–98.

Wang, Y. N., Luo, W. H., Liao, X., & Zhao, P. F. (2024). Exploring the effect of teacher autonomy support on Chinese EFL undergraduates' academic English speaking performance through the mediation of basic psychological needs and classroom engagement. *Frontiers in Psychology*, *15*, 1323713. https://doi.org/10.3389/fpsyg.2024.1323713

Wu, M. L. (2010). Practical questionnaire statistical analysis: SPSS operation and application. *Chongqing University Press*. (In Chinese)

Zachariou, A., & Bonneville-Roussy, A. (2024). The role of autonomy support from teachers in young learners' self-regulation in dyadic contexts: An examination through three-level multilevel analysis. *Learning and Instruction*, *89*, 101843. https://doi.org/10.1016/j.learninstruc.2023.101843

Zhao, F., Yang, Z. S., & Xing, L. H. (2025). The chain mediating role of interest and physical activity level in the PE teacher autonomy support to primary students' physical and mental health. *Scientific Reports*, *15*(1), 37070. https://doi.org/10.1038/s41598-025-21118-y