Psychological Unsafety in Schools: The Development and Validation of a Scale[*]

Kamil Yildirim[1], Şenyurt Yenipinar[1]

[1]Education Faculty, Department of Educational Administration, Aksaray University, Aksaray, Turkey

Correspondence: Kamil Yildirim, Education Faculty B Blok 312, Aksaray University, 068100 Aksaray, Turkey.

Abstract

Work related psychological unsafety defined as the degree to which employee perceive the risky work environment that hinder them to behave comfortably or to speak up what they think without fear of its possible negative consequences. The importance of psychological unsafety has already been laid down for organizational viability and development, but the main problem is lack of instrument to measure psychological unsafety in schools. So the aim of this study is to develop valid and reliable instrument. We developed it in line with the procedures suggested by AERA (The American Educational Research Association). Based on the literature review, we wrote 27 draft items, then we piloted it on 164 school staff and we implemented factor and item analyses. We finally conducted it in the main application with 393 school staff who randomly selected from schools located in different geographical regions of Turkey. Results of EFA displayed that the 19 items loaded on three factors explaining .73 of total variance and the factor loadings ranged between .59 to .85. In addition, corrected item-total correlations ranged from .51 to .86. The internal consistency reliability coefficient was .96 and CFA confirmed the structural model. At the end of the study, we obtained a valid and reliable scale to measure psychological unsafety in schools.

Keywords: psychological unsafety, school, scale development, teachers, administrators

1. Introduction

Employees in organizations interactively work in the direction of defined objectives. Social exchange theory depends mainly on interaction, reciprocity and expectation. According to theory, achieved expectations create the trust among parties in long period (Miles, 2012). Psychological safety, closely related to trust, refers to the environment in which people are comfortable about the consequences of interpersonal risks. It is a belief that one is able to express his or her self without fear of negative consequences to self-image, status or career (Edmondson, Kramer and Cook, 2004; Kahn, 1990). Psychological safety means that you feel comfortable about your colleagues' ideas and actions related with you. Inversely, psychological unsafety indicates more conflict and less team satisfaction among group members. Thus, it is an important aspect of workplace climate (Hernandez, Luthanen, Ramsel and Osatuke, 2015). Workplace climate is a basic condition for psychological safety. Holley and Steiner (2005) suggested discomfort and the presence of risk as the essential components of psychological (un)safety. Psychological unsafety can be regarded as an implicit concept in terms of organizational development. Before occurring much more serious problems, employees must share their ideas, experiences, mistakes and discuss freely without fear of humiliation to prevent them beforehand (Schermerhorn, Hunt, Osborn and Uhl-Bien, 2010). Psychological safety in a team promotes to learn from mistakes and failures (Carmeli and Gittell, 2008). Garvin, Edmondson and Gino (2008) classified the psychological safety as a sub-dimension of supportive learning environment in learning organization concept. It is personal perceptions about how others would respond when you put forward an idea, ask a question, talk about a mistake, or stand out for a position. You can ask yourself "If I do it, will I be hurt, embarrassed or criticized?" (Edmondson et al., 2004). If your responses tend to be "yes" that refers the psychological unsafety. It threats productive discussion that useful for organizational improvement. This situation might prevent the person from fully reflecting his or her potential to work; in turn, it can harm organizational outcomes (Carmeli and Gittell, 2008; Kramer, 1999, Schein, 2010).

---

[*]This study was verbally presented in the 12th International Congress of Educational Administration held in Ankara, May 11-13.

Although psychological safety is closely related to trust, its some characteristics distinguish it from the trust. According to Edmondson et al. (2004) trust is enabler of psychological safety. However, trust does not capture how an employee feels comfortable in work setting. This difference also applies between psychological unsafety and distrust. Trust is a belief made up of reciprocal expectations, but psychological safety belongs to individual. Organizational trust implies that employees should anticipate each other's actions and these actions will have positive consequences for themselves (Miles, 2012). However, psychological safety is an individual matter and it does not require reciprocity. If parties involve in reciprocal safety obligation, we called "psychological contract safety" (Walker, 2010). When we distinguish psychological safety and trust, we consider specifically object of focus (self vs other) (Edmondson et al., 2004).

Definitions and comparisons related with *psychological safety* posit indicators and the main theoretical sub-dimensions of the concept that we can use to measure it. These can be speaking up about errors, failures and problems; feeling comfort for offering new things, criticizing behaviours, situations etc. and having discussion on some matters (Carmeli and Gittell, 2008; Kahn, 1990; Edmondson et al., 2004; Edmondson, Higgins, Singer and Weiner, 2016). However, psychological unsafety, on the other side of the coin, consists interpersonal threats, embarrassment, timidity, exclusion, indifference, lack of support, fear, refrain from explanation etc. According to Carmeli and Gittell (2008) "blaming each other for failures" in a teamwork indicates the psychological unsafety. Conversely, "psychological safety needs to embrace failure as an occasion for learning". If employees have psychological safety, they engage work processes and tasks and they are not outsiders (inclusion vs exclusion) in their organization so they can say what they think and feel (Kahn, 2007). Low psychological unsafety indicates "positive social context in which people feel safe to perform and act" (Carmeli and Gittell, 2008). Williams, Woodon and Wallace (2016) proposed attunement, authenticity and power sharing as facilitators of psychological safety in their study. Garvin et al. (2008) suggested five items measuring psychological safety (e.g. "In this unit, it is easy to speak up about what is on your mind"). Lyu (2016) used three-item scale in examining Chinese teachers' psychological safety. His measurement tool asks whether employees feel comfortable when they express themselves, if there is a threatening climate in the work setting (e.g. "I am not afraid to express my opinions at work." Hernandez, Luthanen, Ramsel and Osatuke (2015) used two items to measure psychological safety in workgroup: "Members in my work-group are able to bring up problems and tough issues" and "It is safe to take risks in this workgroup". Edmondson et al. (2004) studied psychological safety among nurses (N=20-30) from two different hospitals and asked a single survey item (if you make a mistake in this team. is it held against you?). They tried to provide a rough index of team psychological safety. Later, they drew from the interview data a six-item survey variable to assess psychological safety and they used them at the individual level of analysis. They developed a measuring instrument with a gradual work spread over time working with different sectors (health, industry, commercial). They finally proposed a seven-item survey measure requires aggregation of individual data for group level data.

Teachers in a school referred to as "instrumental community". The instrumental community is defined as associations that come together for a specific purpose and manifest themselves through events-interactions (Lunenburg and Ornstein, 2013). Educational institutions work in the form of teamwork, where face-to-face communication dominates in a certain affinity (Karataş, 2009). They interact each other and they share common responsibilities that require them work together. According to Edmondson et al. (2004) in such condition, psychological safety (or unsafety) tend to be experienced. This phenomenon is an implicit condition for schools but it is not fully focused on. The examination of this phenomenon in different cultural context might be important in terms of organizational learning, school improvement and school effectiveness. As noted by Edmondson et al. (2004) there is a need for a valid and reliable measure of psychological safety (or unsafety):

> *...future research must continue to collect data from a variety of team and organizational settings. For example, the effect of psychological safety in culturally diverse teams warrants further research...*(Edmondson et al., 2004: 36).

The main objective of this study is to develop a valid and reliable scale for measuring psychological unsafety in Turkish schools. Literature review did not produce any study focused on this phenomenon yet in Turkish context. Therefore, we would first time measure the phenomenon in Turkish cultural context. In addition, we aimed also improving Edmondson's seven-item scale by depending on different studies carried out at different professional and cultural contexts.

2. Method

This study is a cross-sectional survey, through which researchers describe the perceptions of large number of participants in a short period (Christensen, Johnson and Turner, 2014; Fraenkel, Wallen and Hyun, 2012; Ruane, 2005). According to Edmondson et al. (2004) the level of psychological safety could be inferred from members' spontaneous reports. In this study, we asked participants to what extent they experience the situation mentioned in the items of the scale.

2.1 Participants

The participants of the study were mainly teachers and administrators who, work in primary, lower secondary and secondary level (ISCED 1, 2 and 3). We have two study groups (pilot and main). An in-service training centre provide us to meet participants, as they are easily accessible and representative. There are seven in-service training institutions of the Ministry of National Education (MoNE) located in different provinces of Turkey (EACEA, 2017). Each activity last 3-5 days and each programme might serve between 55 and 100 trainees who come from different schools located in different geographical regions. We collected data through questionnaires administered in "Aksaray In-service Training Institution" between February and April in 2017. Participants consisted of those who are from very different provinces representing all geographical regions of Turkey but they are overwhelmingly from Adıyaman, Aksaray, Bolu, Batman, Düzce, Konya and Istanbul. We delivered questionnaires to 180 participants for pilot study but 164 of it were analysed and data belonged to 393 out of 448 participants for main study were analysed. Table 1 shows information about participants at both applications. There are different opinions about the size of the sample in the course of the scale development process. Nevertheless, minimum size should be five times the number of items in the scale and ideal size should be over 300 (Seçer, 2015). According to this finger account, for a total of 27 items, it is necessary to have minimum 135 participants (five participants per item). Hence, the study satisfies the minimum sample requirement.

When we look at the demographic characteristics of participants at both implementations, the sample might be accepted as representative of the portray of school staff in Turkey (MoNE, 2017). The proportion of male (52-55 %) and female (45-48 %) participants at both applications were very close each other and they were mainly teachers. However, for administrative positions there is an unbalanced proportion against female principals (MoNE, 2017). While the school level increases, inversely the number of participants decreases. A substantial proportion of participants (44%) had at least a four-year experience at the current school that refers participants were very familiar with school environment and school settings. Almost two-thirds of them worked in big schools, which have more than 30 teachers. And all participants have strong job security.

Table 1. Demographic information about participants

Characteristics of Participants		Pilot Study				Main Study			
	f/%	1	2	3	Total	1	2	3	Total
Gender	f	79	85	-	164	176	217	-	393
1: Female; 2: Male	%	*48.0*	*52.0*	*-*	*100*	*45.0*	*55.0*	*-*	*100*
Position	f	15	141	8	164	40	338	15	393
1:Administrator 2:Teacher; 3:Other	%	*10.2*	*86.0*	*3.8*	*100*	*10.2*	*86.0*	*3.8*	*100*
Level	f	69	54	41	164	160	149	84	393
1:Primary; 2:Lower secondary; 3:Secondary	%	*42.1*	*32.9*	*26.0*	*100*	*40.7*	*37.9*	*21.4*	*100*
Experience at current school	f	27	65	72	164	73	139	171	393
1:First year; 2:2-3 year; 3: 4 and more	%	*16.5*	*39.6*	*43.9*	*100*	*18.6*	*35.4*	*46.0*	*100*
School size (Number of teachers)	f	11	40	113	164	41	114	238	393
1:Small; 2:Medium; 3: Big	%	*6.7*	*24.4*	*69.9*	*100*	*10.4*	*29.0*	*60.6*	*100*

2.2 Scale Development Procedures

We followed the standards of scale development procedures recommended by AERA (The American Educational Research Association's Standards for Educational and Psychological Test, 1999) were used. We fulfilled the following steps:

  i)   *Description of the objective of the scale:* Developing a valid and reliable scale to describe the psychological unsafety perceived by teachers and administrators in primary, lower secondary and secondary school levels,

  ii) *Defining the target group:* They are mainly teachers but we focused on school staff including administrators. They are working in compulsory formal education; therefore, private schools are excluded from this study. The proportion of private education institutions at the primary, lower secondary and secondary level are 4.3%, 5.7% and 11.1% accordingly (MoNE, 2017).

  iii) *Examining the other scales:* We reviewed studies focused on psychological safety (Carmeli and Gittell, 2008; Edmondson et al., 2004; Edmondson et al., 2016; Garvin et al., 2008; Kahn, 1990) We especially benefited from Edmondson's (1999) seven-item psychological safety scale. Based on the literature review, we described the theoretical scope of this concept by extracting positive and negative binary items (inclusion-exclusion, attention-indifference, dignify-undignify, tolerance-intolerance, give support-lack of support, facilitation-inhibition, relaxing-making anxiety, encouragement-timidity).

  iv) *Creating item pool:* Based on the literature and structures of the concept, we wrote 27 items consisting both negative and positive meanings,

v) *Drafting a measurement tool:* We produced a draft tool in considering three experts' views (one teacher, two academicians, one of them is expert on assessment and evaluation). It contains 25 items demanding participants to respond the extent to which they experience the statements using five likert type scales (between 1: Definitely represent me and 5: Definitely not represent me).

vi) *Performing pilot study:* We delivered 180 questionnaires and we analysed 164 of it. Information about participants can be seen on Table 1.

vii) *Analysing the data of pilot study:* We eliminated six items from the scale considering exploratory factor analysis. We had 21 items stated in negative meanings under four factors.

viii) Performing main study: We delivered 448 questionnaires and we analysed 393 of it after eliminating improper ones.

Analysing the data of main study and having final version of the scale: We had 19 items under three factors (exclusion, lack of support and timidity) depending on the results of item analysis, exploratory and confirmatory factor analysis.

2.3 Collecting Data

Pilot study was implemented in February-March period; the main application was performed in March-April period. In both applications, we worked with the school staff covering administrators and teachers who participated in trainings in Aksaray In-Service Training Institute of MoNE. The scale was administered together with a demographic questionnaire consisting of 10 items. Permissions were obtained from relevant institutions before the application of the scales. When the participants were present together, we first informed them about the scale and delivered to the volunteer participants to fill out the questionnaire forms including the scale. Completing the questionnaires lasted approximately 20-25 minutes.

2.4 Analysing Data

Before implementing the analysis, we prepared the data for analysis. We first examined questionnaires whether they were properly filled. After eliminating improper forms, we coded each questionnaire, and then we entered the raw data into computer. We checked the data in terms of outliers, duplications and missing. We performed descriptive analysis techniques (f, SD), Pearson Correlation Coefficient (r), exploratory factor analysis (varimax rotation) and confirmatory factor analysis. We performed all analysis at individual level. Analysis were carried out using SPSS version 22.0 and LISREL 8.7. When we interpreted the continuous data, which refers to the level of experiencing psychological unsafety, we regarded the following classification: 1.00-1.80=Very high, 1.81-2.60=High, 2.61-3.40=Moderate, 3.41-4.20=Low, 4.21-5.00=Very low.

3. Results

3.1 Results of Pilot Test

We first eliminated improper questionnaires then we checked if there were outliers or duplications. We identified six missing data using descriptive statistics and replaced them with series mean. We computed mahalanobis distances to reveal outliers and we cleared only one cases because their scores exceeded the limit (Seçer, 2015; Tabachnick and Fidell, 2013). There were no duplications.

First version of the scale had 24 items but experts suggested one items (*If I have an opportunity, I want to go another school because of psychological unsafety*) to add into the scale therefore, psychological unsafety scale had 25 negatively stated items at this stage. We analysed the data belonged to 164 participants in pilot study. Table 2 shows descriptive information of participants' responses. Mean values of items range from 3.16 to 3.63 up to 5.00 that indicate the matters stated in the scale are experienced moderate or low level. There is no statement informed happening at neither very high or very low levels. General mean score for the overall scale is x̄=3.42 indicating that school staff experience the psychological unsafety at low level but it is slightly over the range of moderate level. Therefore, psychological unsafety is not negligible. Mean scores of 12 items are lower than general mean scores that they are experienced at moderate level. Comparatively the most negative statement (PS1, x̄=3.16) is "*If you make a mistake in this institution that is used against you later*". Some of other negative statements (PS13, x̄=3.27) are "*Although I want to talk about some matters but I stay silent because of its possible negative consequences*" and (PS2, x̄=3.32) are "*Anybody demands someone else's help unless s/he has to do*".

Comparatively the most positive statements are (PS18, x̄=3.63) "*I feel that people keep important information to themselves and do not share it with others.*" Other positive statements are (PS5, x̄=3.55) "*I feel that people in this school do not take me acceptance.*" and (PS7, x̄=3.55) are "*I do not count on our administrators to defend me in any case*". Based on the perceptions of participants, initial result refers that psychological unsafety matters happen at low level in Turkish school.

Table 2. Descriptive statistics results of pilot test implementation

Code	İtems		SD
PS1	If you make a mistake in this institution, that is used against you later.	3.16	1.41
PS2	Anybody demands someone else's help unless s/he has to do.	3.32	1.28
PS3	In this institution, anybody encourage someone else for improvement.	3.43	1.28
PS4	Anybody support someone else for getting better position and tasks.	3.43	1.35
PS5	If I have an opportunity, I want to go another school because of psychological unsafety.	3.35	1.49
PS6	I could not state any matter to administrators.	3.38	1.57
PS7	I do not count on our administrators to defend me in any case.	3.55	1.49
PS8	I do not count on my colleagues to defend me in any case.	3.35	1.51
PS9	When I submit a problem to administrators, they remain irrelevant.	3.52	1.43
PS10	When I submit a problem to my colleagues, they remain irrelevant.	3.35	1.38
PS11	People in this school cannot easily bring controversial issues.	3.36	1.40
PS12	In situations that contradict me, I hesitate to appeal.	3.43	1.51
PS13	I want to talk about some matters but I stay silent because of its possible negative consequences.	3.27	1.38
PS14	I have never been involved in issues that could have negative consequences for me.	3.34	1.42
PS15	There may be deliberate attempts at this institution to dismiss my efforts.	3.49	1.50
PS16	There may be behaviours in this place that may diminish the value of my work.	3.35	1.45
PS17	I feel that my colleagues, who I work with, do not value my knowledge and skills.	3.48	1.48
PS18	I feel that people keep important information to themselves and do not share it with others.	3.63	1.34
PS19	I think that people in this place share important information only with those close to them.	3.37	1.38
PS20	I do not feel "we" feelings in this place.	3.32	1.45
PS21	I feel that no one will take me in any way.	3.53	1.54
PS22	I think that people in this place can apply for unethical ways to come into prominence.	3.51	1.43
PS23	I feel that people in this place do not accept each other as they are.	3.49	1.34
PS24	I feel that I am excluded from the network of information, experience and material exchange.	3.54	1.45
PS25	I feel that people in this school do not take me acceptance.	3.55	1.62
Total		3.42	1.44

We performed exploratory factor analysis (EFA) on 25 items for data reduction. We first checked the appropriateness of the data for EFA. As they met the criteria of EFA (KMO=0.96, Bartlett's Test of Sphericity=3726.49 and p=0.000) we conducted it. We selected principal component analysis, opted to analyse correlation matrix, chose eigen value over one as extraction technique and we applied varimax rotation, then we suppressed absolute values less than .30. After the first application, we had four factors that explained .72 of total variance. Because all variables sorted by size and grouped, we easily determined which variable is associated with which factor. When we examined the rotated solution, we noticed four items (PS5, PS6, PS19 and PS20), which have loading values under different factors are less than .10. These variables are obstacles for separating each factor from others (Pituch and Stevens, 2016) therefore; we removed them from analysis in accordance with stepwise technique. We finally had 21 items under four factors explaining .75 of total variance. Contributions of each factor in explaining total variance are F_1=.33, F_2=.16, F_3=.14 and F_4=.12. Then we conducted internal reliability test for these 21 items. Internal consistency of the scale is the more common index of reliability. Cronbach's Alpha (α) was .96 that indicates high internal consistency of the scale. The Item-Total statistics displays information on how individual item contribute to the whole. All items positively contributed to the construct the scale measures. Item-total correlations ranged from .39 to .82. They suggested that these items are valid because of no items had value less than .30 (Gerber and Finn, 2005). When we looked at internal consistency for sub-scales, we noticed that they had high internal consistency scores. *Exclusion* had nine items and its α is .96; *indifference* had five items and its α is .90; *lack of support* had four items and its α is .82 and finally *timidity* had three items and α is .85. Based on the preliminary results psychological unsafety scale was valid and reliable. Depending on the pilot study, we further stepped for performing the main study.

3.2 Results of Main Test

After conducting questionnaire including psychological unsafety measure at the main study, we prepared the data for analysis. We first eliminated 15 improperly filled questionnaires (e.g. much of them is blank or same coded for all items) out of total 448 questionnaires. Then we coded proper questionnaires and transformed data into SPSS software programme. Checking for outliers did not provide any case however; test for duplications showed 40 duplications therefore we removed them from analysis. We finally analysed data belonged to 393 participants in the main application. Table 3 shows the contents, factor loadings and results of item analysis.

Table 3. Results of factorial constructs and item analysis (N=393)

Factor FactorItems Loadings	Std. Mean Dev.	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
PS24 .838	3.63 1.44	63.22	371.35	.827	.797	.958
PS15 .838	3.57 1.50	63.28	370.48	.810	.845	.958
PS16 .835	3.51 1.47	63.34	370.80	.818	.849	.958
PS25 .798	3.70 1.54	63.15	367.89	.830	.801	.958
PS17 .782	3.58 1.46	63.27	370.93	.821	.748	.958
PS21 .781	3.67 1.45	63.18	370.42	.837	.766	.957
PS22 .763	3.56 1.44	63.29	370.64	.839	.745	.957
PS9 .730	3.59 1.40	63.26	375.80	.764	.720	.959
PS23 .726	3.59 1.32	63.26	375.37	.825	.722	.958
PS18 .709	3.61 1.39	63.24	373.62	.814	.718	.958
PS10 .617	3.51 1.35	63.35	378.26	.747	.654	.959
PS7 .593	3.49 1.45	63.37	376.44	.723	.600	.959
F_1 Averages	3.59 1.44	63.27	372.67	.807	.754	.958
PS2 .802	3.35 1.28	63.50	390.06	.548	.492	.961
PS3 .751	3.45 1.31	63.41	381.98	.699	.673	.959
PS4 .735	3.41 1.36	63.44	379.33	.722	.694	.959
PS1 .698	3.19 1.37	63.66	392.31	.462	.325	.963
F_2 Averages	3.35 1.33	63.50	385.92	.617	.552	.961
PS14 .845	3.38 1.39	63.47	386.87	.557	.549	.961
PS13 .841	3.48 1.33	63.37	385.22	.619	.640	.960
PS12 .671	3.57 1.41	63.28	378.86	.702	.634	.959
F_3 Averages	3.48 1.38	63.37	383.65	.626	.608	.963
Scale Averages	3.52 1.40	63.33	377.19	.735	.683	.959

(Factors: F_1: Exclusion, F_2: Lack of Support, F_3: Timidity)

We implemented EFA to test construct validity of the measure. Before implementing it, we checked the conformity of the data for analysis. KMO (.95), Bartlett Test of Sphericity (6749.03) and X^2 is significant (p=.000) refer to its conformity to be analysed. We selected the same methods and techniques as we did in pilot stage (Principal component, correlation matrix, eigen value over one, varimax rotation, suppressed less than .30). First application showed that 21 items go under three factors. Total variance explained is .71. Examining the rotated solution indicated that there are two variables (PS8 and PS11), of which value of factor loadings at different factors are smaller than .10, so that we performed analysis without these variables in terms of stepwise technique. We had 19 items under three factors explaining .73 of total variance. Contributions of each factor in explaining total variance are F_1=.40, F_2=.18 and F_3=.14. Factor loadings of items ranged from .59 to .85. Results of EFA indicate that validity of the measure is at very good level (Gerber and Finn, 2005; Pituch and Stevens, 2016).

We named the factor considering its contents as F_1: *Exclusion,* F_2: *Lack of support* and F_3: *Timidity*. We examined contribution of each item to the overall internal consistency of measurement by scores belonged to item-total correlations, squared multiple correlations and Alpha if item deleted. We considered the scores smaller than .30 as criteria for deletion. For both corrected item-total correlation coefficients and mean scores there is no score smaller than the criteria. Therefore, we did not remove any item from the measurement and at the conclusion of the item analysis, psychological safety scale comprised 19 items and Cronbach's Alpha indicated high internal consistency (α=.964) of the scale. We computed the internal consistency for each sub-scale and scores indicated that they are highly reliable (αF_1=.96, αF_2=.84 and αF_3=.85).

The mean for overall scale is 3.52 that indicate the existence of low level of psychological unsafety in Turkish schools. However, SD (1.4) refers considerable differences among participants in terms of psychological unsafety perceptions. When we examined the factor averages, we determined that participants experience *lack of support* at moderate level (x̄=3.35. SD=1.33) that has comparatively remarkable result. Participants perceived that they moderately experience lack of support. They experienced *timidity* at low level (x̄=3.48. SD=1.38). The most positive situation in terms of psychological unsafety is that participants experience relatively the lowest level of *exclusion* (x̄=3.59, SD=1.44).

Confirmatory factor analysis (CFA) is commonly used during the scale development process (Çokluk, Şekercioğlu and Büyüköztürk, 2010; Kline, 2011; Pituch and Stevens, 2016). We need to verify whether the constructs attract their items (observed variables) as we defined by EFA. Therefore, we carried out CFA. When implementing CFA, we selected the maximum likelihood estimation method and covariance matrix on LISREL 8.7 software programme. We evaluated the goodness of fit by the χ^2 test and the RMSEA (root mean square error of approximation), CFI (comparative fit index) and NNFI (non-normed fit index). We regarded goodness of fit indices (χ^2/df≤5, p=0.000, .05≤RMSEA≤.10, .95≤CFI≤.97, .95≤NFI≤.97, .90≤GFI≤.95) in understanding whether the results verify constructs (Çokluk et al., 2010; Kline, 2011). Table 4 shows the goodness of fit statistics.

Table 4. Goodness of fit statistic results

Model	χ^2	$(\chi^2/df)^*$	RMSEA	NFI	NNFI	CFI	GFI	AGFI
PuS	742.80	4.90	0.10	0.97	0.97	0.97	0.83	0.79

*df = 149, p < 0.01

Checking t values displayed significant of observed variables at the .01 level. Therefore, we did not remove any observed variable. Significance of the χ^2 informs us that the model fit with the data and p value verified this result. Other fit indices verified the constructs because NFI, NNFI and CFI exceed .90 which is recommended cut-off for these indices. The value of RMSEA (.10) refers a mediocre fit (Kline, 2011; Pituch and Stevens, 2016).

According to Kline (2011) "A path model is a structural model for observed variables and this structural model represents hypotheses about the effect priority". EFA proposed the hypotheses thus we can see whether they are verified through a reverse transaction that we test the observed variables go under the hypothesised constructs. This application requires *standard CFA model* in which we can see direct effects between factors and indicators. We performed second order path analysis, thus Figure 1 shows the links between factors (exogenous variables) and indicators (observed variables or endogenous variables). As we noticed on Figure 1, path analysis verified the constructs that we had through EFA (Çelik and Yılmaz, 2016; Kline, 2011). Since the results refer an acceptable (mediocre) fit, we did not make any corrections to get better fit indices because the constructs were confirmed already. Results of EFA and CFA together with item analysis validated the 19 item-work related psychological unsafety scale in Turkish schools.
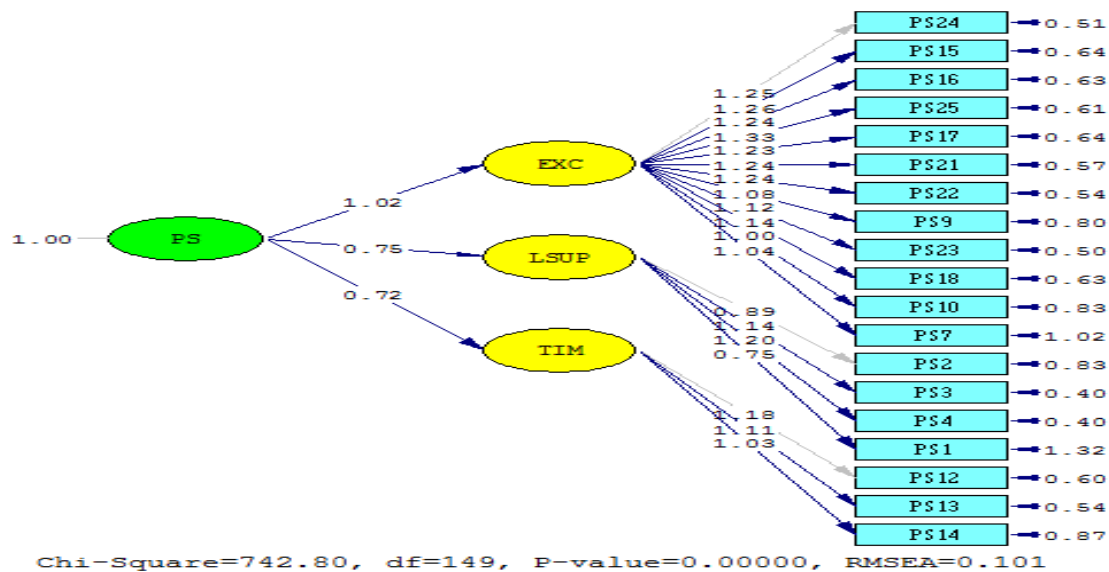

Figure 1. Path analysis of psychological unsafety scale constructs

4. Discussion

Because of its importance for organizational viability and development, psychological unsafety requires investigations. However, scarce number of studies focused on this phenomenon appeared an obstacle, specifically in different cultural context. Starting point in setting off this research is to have a valid and reliable measurement tool. This study originated from such a need. Thus, developing *the work related psychological unsafety scale* helps reducing the scarcity of research focused on this particular topic. It is not only a basis for further studies in this area but also presents the situation in culturally different context.

In this study, we developed and validated the work related psychological unsafety scale for Turkish schools. During the scale development process, we followed standards of AERA (1999). Based on the literature review, we wrote 27 items consisting positive and negative statements. After taking experts' views, we made corrections and we piloted the draft scale consisting all negatively stated 25 items. After refinement on the draft scale, then we implemented it on totally 393 school staff randomly selected from schools representing seven geographical regions of Turkey. Based on the EFA, item analysis and CFA, we had a validated the scale measuring *the work related psychological unsafety* in Turkish schools. It consisted of 19 negatively stated items scaled with Likert type ranging from 1: very high to 5: very low. Higher score means low psychological unsafety (or high psychological safety). The scale had good psychometric properties. EFA produced three factors and CFA verified these constructs.

When we compare its psychometric properties (α=.96) to those of previous studies, it can be said that we have had now better measurement tool. Internal consistency of Edmondson's (1999) scale was .70 and Carmeli and Gittel (2008) used later Edmondson's scale and computed its internal consistency as .84. In addition, Lyu's (2016) three-item scale had .54 internal consistency score. We also improved Edmondson's seven-item scale in different cultural context. As the participants comprised primary, lower secondary and secondary schools in this study that also distinguished it from studies carried out at only one school level (e.g. Lyu, 2016) and only particular staff (e.g. Hjerto and Paulsen, 2017).

As a result, we found that school staff experience the work related psychological unsafety at low level in Turkey. Among the sub-scales, the most notable one is the *need of support* in school settings. Participants experience *lack of support* at moderate level. The other sub-scales, *exclusion* and *timidity*, are experienced at low level. Although results indicate that school staff experience the psychological unsafety at low level but if we consider its mean, which is close to moderate level, we can say that psychological unsafety is not negligible. Furthermore, standard deviations indicate big variances among participants in terms of experiencing psychological unsafety. Some previous research findings related with Turkish teachers are coherence with the results of the current study that teachers experience some degree of psychological unsafety. When we count on the result suggesting "psychological safety is an important aspect of workplace climate" (Hernandez et al., 2015) we can relate the current results to the results of school culture, school climate and professional well-being. For instance, Yaman and Güngör (2014) found that teachers had a low level of exclusion bias. Sezgin, Koşar and Er (2014) qualitatively examined the mentoring process in schools and found that while school principals tend to support their staff but enriching the co-operative works is not adequate in terms of professional development. Uğurlu and Abdürrezzak investigated whether teachers consider their principal as leader and their study resulted in that teachers' the most important expectations are encouragements, co-operation efforts and empathizing ability in personal relations when they need. In addition, collaboration among teachers, supportive culture and recognition are found being lower than expectation (Demirtaş, 2010; Terzi, 2005; Yıldırım, 2015). This coherence can be considered as supporting external validity of the current scale.

The current study's findings largely concur with that of previous studies carried out at different cultural context. Small differences may originate from participants' characteristics. While mean scores are more compatible but the variances are quite different particularly for the current study. When reader evaluate the findings of following studies discussed in this paragraph should consider that items are measured on a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree). e.g. Carmeli and Gittell (2008) studied on employees who work in the software, electronics, and finance industries found that medium level of psychological safety ($\bar{x}$=3.44, SD=.68) that refers medium psychological unsafety. Hjerto and Paulsen (2017) examined the place of psychological safety in complex relationships among individual team learning on 171 school principals and deputy principals who are together responsible in managing school. They collected the data by internet based survey method. They found high level of psychological safety ($\bar{x}$=4.15, SD=.61). Lyu (2016) examined psychological safety based on the perceptions of totally 254 Chinese teachers working in compulsory schools and found that they have medium level of psychological safety ($\bar{x}$=3.42, SD=.78). Similar results ($\bar{x}$=3.42, SD=.52) were found by Hernandez et al. (2015) on 681 Veteran's Health Administration supervisors in locations across the United States. They gathered data from a five-year period internet based national studies. Williams et al. (2016), implementing a qualitative study, examined the theoretical ingredients of psychological safety in terms of a controversial topic (racial dialogue) on purposefully selected two classrooms in USA. They found teachers' attitudes and behaviours regarding attunement, authenticity and power sharing would develop students' psychological safety. Mean scores also change with the participants' different demographic characteristics such as gender, experience, tenure etc. For instance, Edmondson et al. (2016) found lower level of psychological safety of novice teachers comparing with their late-career colleagues.

Results of the studies focused on psychological (un)safety, which were carried out with different professional groups and in different cultural contexts, suggested a coherence with the results of current study. Thus, we argue that external validity of the scale for different cultural context at good level and it would be used for them, too.

Given the "Edmondson's theory that psychological safety is necessary for enabling effective learning in organizations", administrators can benefit this study's results in order to enhancing organizational learning. Taken together the results of Carmeli and Gittell (2008), Hernandez et al. (2015) and of this study, administrators should encourage school staff to explain their opinions and share failures instead of covering it up. Administrators can use leadership roles in forming psychologically safe environments in school (Hernandez et al., 2015). School administrators can also use the scale and portray the level of psychological unsafety; in addition, they can benefit the results of this study in developing their organization in terms of *exclusion*, *lack of support* and *timidity*.

Participants of this study may not represent population because the study excluded private schools and their staff therefore; the scale developed in this study should be used with different samples in order to test its reliability. Sampling method and sampling size may constraint the generalizability of the results. Furthermore, cultural context should be

considered in generalizing the results of the study.

At the end of the study we developed valid and reliable scale measuring psychological (un)safety in schools in Turkish cultural context. We also revealed that school staff including teachers, principals and deputy principals experience low level of psychological unsafety.

References

AERA (1999). *Standards for educational and psychological testing.* Washington. DC: American Educational Research Association.

Carmeli, A., & Gittell, J. H. (2008). High-quality relationships, psychological safety, and learning from failures in work organizations. *Journal of Organizational Behavior, 30*, 709–729. https://doi.org/10.1002/job.v30:6

Çelik, H. E., & Yılmaz, V. (2016). *Yapısal Eşitlik Modellemesi-Temel Kavramlar-Uygulamalar-Programlama [Structural Equation Modeling-Basic concepts-applications-programming]*. Ankara: Anı

Christensen, L. B., Johnson, R. B., & Turner, A. L. (2014). *Research methods design and analysis*. New York. Pearson Publishing.

Çokluk, Ö., Şekercioğlu, G., & Büyüköztürk, Ş. (2010). *Statistics for multiple variables: Practices in SPSS and LISREL.* Ankara: PegemA.

Demirtaş. Z. (2010). The relationship between school culture and students' performance. *Education and Science, 35*(158), 1-13.

EACEA (Education Audiovisual & Culture Executive Agency). (2017). *Continuing Professional development for teachers working in early childhood and school education.* Retrieved from https://webgate.ec.europa.eu on 17.03.2017.

Edmondson, A. C. (1999). Psychological safety and learning behaviour in work teams. *Administrative Science Quarterly, 44*, 350-383. https://doi.org/10.2307/2666999

Edmondson, A. C., Higgins, M., Singer, S., & Weiner, J. (2016). Understanding Psychological Safety in Health Care and Education Organizations: A Comparative Perspective. *Research in Human Development, 13*(1). 65-83. https://doi.org/10.1080/15427609.2016.1141280

Edmondson, A. C., Kramer, R. M., & Cook, K. S. (2004). Psychological safety. trust. and learning in organizations: A group-level lens. *Trust and distrust in organizations: Dilemmas and approaches, 12*, 239-272. Retrieved from https://www.researchgate.net on 28.01.2017

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education.* New York: McGraw Hill.

Garvin, D. A., Edmondson, A. C., & Gino, F. (2008). Is yours a learning organization? *Harvard Business Review, 86*(3), 109-116.

Gerber, S. B., & Finn, K. V. (2005). *Using SPSS for windows: data analysis and graphics*. NY: Springer.

Hernandez, W., Luthanen, A., Ramsel, D., & Osatuke, K. (2015). The mediating relationship of self-awareness on supervisor burnout and workgroup civility and psychological safety: a multilevel path analysis. *Burnout Research, 2*, 36-49. https://doi.org/10.1016/j.burn.2015.02.002

Hjerto, K. B., & Paulsen, M. (2017). Learning outcomes in leadership teams: The multilevel dynamics of mastery goal orientation, team psychological safety, and team potency. *Human Performance, 30*(1), 38-56, https://doi.org/10.1080/08959285.2016.1250765

Holley, L. C., & Steiner, S. (2005). Safe space: Student perspectives on classroom environment. *Journal of Social Work Education*, *41*(1), 49–64. https://doi.org/10.5175/JSWE.2005.200300343

Kahn, W. A. (1990). Psychological conditions of personal engagement and disengagement at work. *Academy of Management Journal, 33*(4), 692-724. https://doi.org/10.2307/256287

Kahn, W. A. (2007). Meaningful connections: Positive relationships and attachments at work. In J. E. Dutton, & B. R. Ragins (Eds.), *Exploring positive relationships at work: Building a theoretical and research foundation* (189-206). Mahwah, NJ: Lawrence Erlbaum Associates

Karataş, E. (2009). *Job values and organizational culture in educational administration.* Yayınlanmamış yüksek lisans tezi [Unpublished master's thesis]. Uşak University Social Sciences Institutions, Uşak.

Kline, R. B. (2011). *Principles and practice of structural equation modeling.* NY: Guilford

Kramer. R. M. (1999). Trust and distrust in organizations: Emerging perspectives. enduring questions. *Annual Review of Psychology, 50*, 569-598. https://doi.org/10.1146/annurev.psych.50.1.569

Lunenburg, F., C&Ornstein, A. C. (2013). *Educational Administration.* (Tran. Ed. G. Arastaman). Ankara: Nobel Yayıncılık. (Original edition 6. Publication year, 2012).

Lyu, X. (2016). Effect of organizational justice on work engagement with psychological safety as a mediator: evidence from china. *Social Behaviour and Personality, 44*(8), 1359-1370. https://doi.org/10.2224/sbp.2016.44.8.1359

Miles, J. A. (2012). *Management and Organization Theories.* San Francisko: Wiley.

MoNE (Ministry of National Education). (2017). National education statistics, formal education: 2015/'16. retrieved fromhttp://sgb.meb.gov.tr/meb_iys_dosyalar/2016_03/30044345_meb_istatistikleri_orgun_egitim_2015_2016.pdf

Pituch, K. A., & Stevens, J. P. (2016). *Applied multivariate statistics for the social sciences.* NY: Routledge.

Ruane, J. M. (2005). *Essentials of research method: A guide to social science research.* USA: Blackwell.

Schein, E. H. (2010). *Organizational culture and leadership.* San Francisco: Jossey-Bass.

Schermerhorn, J. R., Hunt, J. G., Osborn, R. N., & Uhl-Bien, M. (2010). *Organizational Behavior.* San Francisko: Wiley.

Seçer, İ. (2015). *The process of development and adaptation measurement: applications of SPSS and LISREL.* Ankara: Anı

Sezgin, F., Koşar, S., & Er, E. (2014). Examining mentoring in the training of school administrators and teachers. *Kastamonu Education Journal, 22*(3), 1337-1356.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics.* Sixth edition. New York: Pearson.

Terzi. A. R. (2005). Organizational culture in elementary schools. *Educational Administration: Theory and Practice. 43.* 423-442.

Uğurlu, C. T., & Abdürrezzak, S. (2016). Research of teachers' views on effective school and school leadership. Educational Administration, *Theory and Practice, 22*(3), 401-428. https://doi.org/10.14527/kuey.2016.016

Walker. A. (2010). The development and validation of a psychological contract of safety scale. *Journal of Safety Research, 41,* 315-321. https://doi.org/10.1016/j.jsr.2010.06.002

Williams, J. D., Woodson, A. N., & Wallace, T. L. (2016). Can we say the N-word? Exploring psychological safety during race talk. *Research in Human Development, 13*(1), 15-31. https://doi.org/10.1080/15427609.2016.1141279

Yaman, E., & Güngör, H. (2014). Teachers' views on stigma tendency. *Adıyaman Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 18*, 823-851. https://doi.org/10.14520/adyusbd.780

Yıldırım, K. (2015). Testing the main determinants of teachers' professional well-being by using a mixed method. *Teacher Development, 19*(1)*,* 59-78. https://doi.org/10.1080/13664530.2014.970663