Instrument (scale) development considerations: A case study of the Authoritative School Climate Survey (ASCS)

2025.11.14

Francis L. Huang / huangf@missouri.edu
Professor / Statistics, Measurement, and Evaluation in Education Program
Educational, School, and Counseling Psychology, University of Missouri
https://francish.net
@flhuang

College of Education & Human Development
University of Missouri

Presentation @Education University of Hong Kong

Abstract

In psychology and education, researchers are often interested in instruments that measure unobserved, latent constructs referred to as scales. Through this talk, Dr. Huang will outline the development and testing of the Authoritative School Climate Survey (ASCS) which has been used by thousands of secondary school students, teachers, and principals. He will discuss the practical aspects of large-scale survey development, design, and administration. Statistical analyses such as (multilevel) exploratory and confirmatory factor analyses will be discussed as well as topics on reliability, validity, and invariance testing.

Agenda

- Conceptual and theoretical background: School climate
- Survey development and administration
- Assessing psychometric properties
- Refinement and testing

School climate has been associated with many positive outcomes

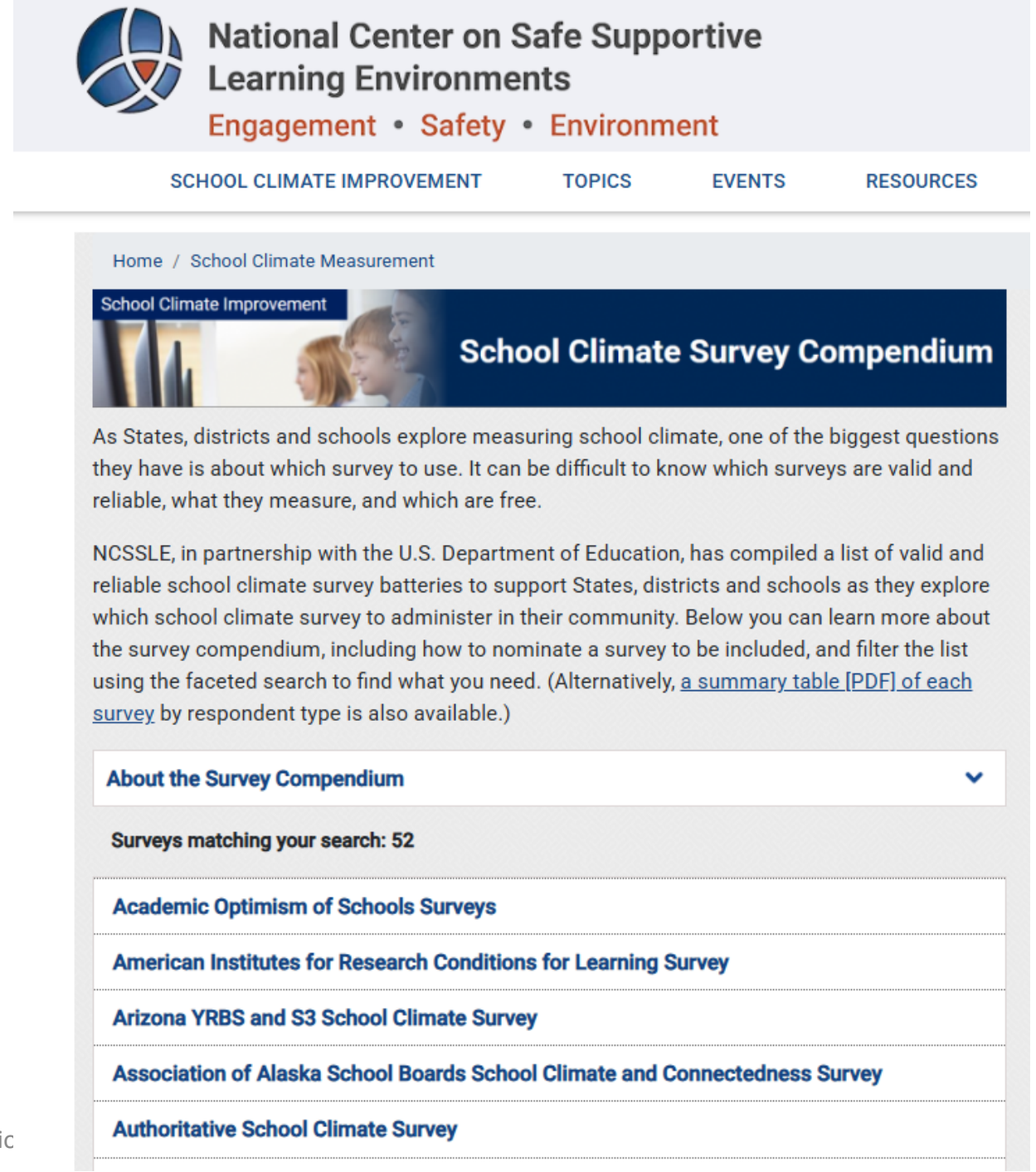
- Improved graduation rates (Thapa et al., 2013)
- Higher engagement and academic achievement (Konold et al., 2018a; Voight & Hanson, 2017)
- Better socio-emotional health (Wong et al., 2021)
- Lower student risk behavior (Cornell & Huang, 2016)
- Reduced use of school suspensions (Huang & Cornell, 2018)

Several definitions of school climate…

- Most widely used: “Quality and character of school life” … [school climate] “is based on the patterns of people's experiences of school life and reflects norms, goals, values, interpersonal relationships, teaching and learning practices, and organizational structures” (Cohen et al., 2009)
- “The quality and consistency of interpersonal interactions within the school community that influence children’s cognitive, social and psychological development” (Haynes et al., 1997, p. 332)

There are dozens of school climate surveys available

- See https://safesupportivelearning.ed.gov/topic-research/school-climate-measurement/school-climate-survey-compendium



School climate should be…

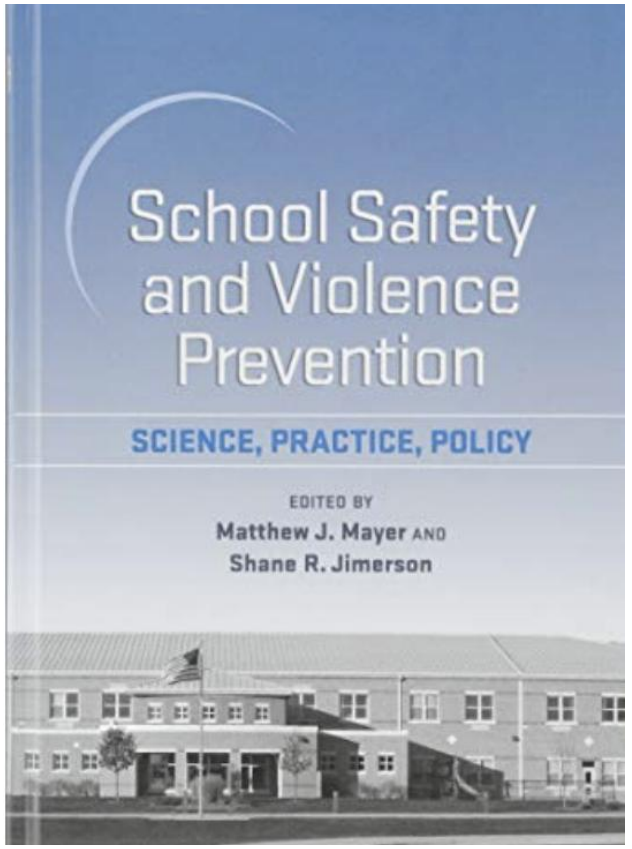
- Measurable (what are some ways?)
- More than a laundry list
- Meaningfully related to one another
- Related to some other student and school outcomes

However, if measured without an overarching framework, school climate encompasses…

- …“just about every feature of the school environment” (Wang & Degol, 2016, p. 317)
- Becomes difficult to disentangle from other school characteristics (Rudasill, Snyder, Levinson, & Adelson, 2018)

If every aspect of the school is part of climate…

- …then it is not clear what the concept means and what are its malleable components (Cornell & Huang, 2017)



Cornell, D., & **Huang, F**. (2019). Collecting and analyzing local school safety and climate data. In M. Mayer & S. Jimerson (Eds.), *School safety and violence prevention: Science, practice, and policy driving change.* (pp. 151–175). Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/0000106-007

Authoritative school climate (ASC) theory (Cornell & Huang, 2016; Gregory et al., 2010) is a conceptual model based on

- Baumrind's (1968) parenting typology and authoritative parenting research

	Low Demandingness (Low Expectations/Control)	**High Demandingness (High Expectations/Control)**
High Responsiveness (High Warmth/Support)	PERMISSIVE (Indulgent)	AUTHORITATIVE
Low Responsiveness (Low Warmth/Support)	UNINVOLVED (Neglectful)	AUTHORITARIAN

- Many adaptations have been made to this basic typology

In a school context, these can be characterized in dimensions of disciplinary structure and student support

- Caveat: As a conceptual model, schooling and parenting are different processes

	Low Structure (Low Expectations/Control)	**High Structure (High Expectations/Control)**
High Support (High Warmth/Support)	PERMISSIVE (Indulgent)	AUTHORITATIVE
Low Support (Low Warmth/Support)	UNINVOLVED (Neglectful)	AUTHORITARIAN

In a school context, these can be characterized in dimensions of disciplinary structure and student support (cont.)

- Disciplinary structure refers to strict but fair enforcement of school rules
- Student support refers to the student experience of teachers and other school staff members as supportive, respectful, and willing to help

	Low Structure (Low Expectations/Control)	High Structure (High Expectations/Control)
High Support (High Warmth/Support)	PERMISSIVE (Indulgent)	AUTHORITATIVE
Low Support (Low Warmth/Support)	UNINVOLVED (Neglectful)	AUTHORITARIAN

Several studies use the authoritative parenting typology in school climate research…

Referred to as ASC theory:

- Fisher et al., 2017
- Gregory et al., 2010
- Lee, 2012
- Pellerin, 2005

	Low Structure (Low Expectations/Control)	High Structure (High Expectations/Control)
High Support (High Warmth/Support)	PERMISSIVE (Indulgent)	AUTHORITATIVE
Low Support (Low Warmth/Support)	UNINVOLVED (Neglectful)	AUTHORITARIAN

- Many other surveys measure these aspects (e.g., rules are fair, positive relationship with teachers) but do no explicitly use ASC (e.g., Bear, Gaskins, Blank, & Chen, 2011; Brand, Felner, Shim, Seitsinger, & Dumas, 2003)

Agenda

- Conceptual and theoretical background: School climate
- Survey development and administration
- Assessing psychometric properties
- Refinement and testing

Prior studies informed what items to initially include in the ASCS

- Many items derived form an earlier survey, the School Climate Bullying Survey (SCBS)
- Complete ASCS had ~100 items consisting of four primary scales, some supplemental scales, and experimental items (and demographic info)
- Response options (for scale items) ranged from: Strongly Agree, Agree, Disagree, Strongly Disagree

Project funding and support were crucial

- Development of a Standard Model for School Safety Assessment (National Institute of Justice, Department of Justice) (2012-2016)
- Improvement of School Climate Assessment in Virginia Secondary Schools, National Institute of Justice (2018-2020)
- Done in coordination with the:
  - Virginia Department of Criminal Justice Services
  - Virginia Department of Education
- See https://education.virginia.edu/research-initiatives/research-centers-labs/research-labs/youth-violence-project/bullying-school-climate/authoritative-school-climate-survey
- Final report: https://education.virginia.edu/documents/development-standard-model-school-climate-and-safety-assessment-final-report

Many practical considerations that are critical…

- Getting schools to participate?
- How to follow up (when you are working with hundreds of schools, can be an issue)
- How to get the online survey system to do what you want (e.g., randomization; gift cards)

Administered to Virginia public students, staff, and principals in the spring of each year (from 2013 – 2020)

Had several thousand student and teacher responses, for example:

Year	Samples (Final)	Grade
2013	39,364 students from 423 schools	7-8
2014	48,027 students from 323 schools	9-12
2015	85,762 students from 410 schools	6-8 (expanded)

For teachers and staff, ~10,000 responses from over 300/400+ schools– had a similar survey. Also had a principal survey

Information available at
https://www.icpsr.umich.edu/web/NACJD/studies/38022

At the end of the SY, schools received a customized report comparing their school, region, and state-level averages

- Had very high school response rates (often > 95% of schools)
- Excluded schools that served a specialized population (e.g., juvenile justice centers, adult learners, students with special needs)
- Needed a report generation system

Student and Teacher/Staff Perceptions of School Climate

Bayside High

Grades 9-12, Spring 2016

The Virginia Secondary School Climate Survey provides schools with a biennial assessment of school climate and safety conditions from the perspective of students and teachers/staff. The purpose of this report is to help schools identify strengths and weaknesses that can guide efforts to improve school safety and student learning.

This report is based on responses from 156 students and 51 teachers/staff in your school. State results are based on 62,679 students and 14,619 teachers/staff in 320 schools, with additional comparisons to schools in your region. A breakdown of student answers by grade and gender is found in a supplementary file available with this report. For more information, see http://www.dcjs.virginia.gov/vcss/audit/student/

Contents

Administered using Qualtrics (not yet very common back then)-- some other considerations--

As a statewide survey:

- Gave administrators an option to use random sampling or whole grade options
  - Required weighting responses later on
- Required constant follow ups which needed constant monitoring/updates
- Had to provide options for randomized test-retest options later on (and provide gift cards)
- Forced choice vs optional?

Administered using Qualtrics-- some other considerations (cont.)--

~6% may be invalid. Degrades survey quality by including 'junk' responses

- Include validity screening items! (many ways)
  - I am telling the truth on this survey (SD, D, A, SA)
  - How many questions did you answer truthfully? (all of them, all but 1 or 2 of them, most of them, some of them, only a few or none of them)
- For staff, "I am paying attention"

Article

The Impact of Validity Screening on Associations Between Self-Reports of Bullying Victimization and Student Outcomes

Educational and Psychological Measurement
2018, Vol. 78(1) 80–102

DOI: 10.1177/0013164416671767
journals.sagepub.com/home/epm
SAGE

Yuane Jia[1], Timothy R. Konold[1], Dewey Cornell[1], and Francis Huang[2]

Psychological Assessment
2012, Vol. 24, No. 1, 21–35

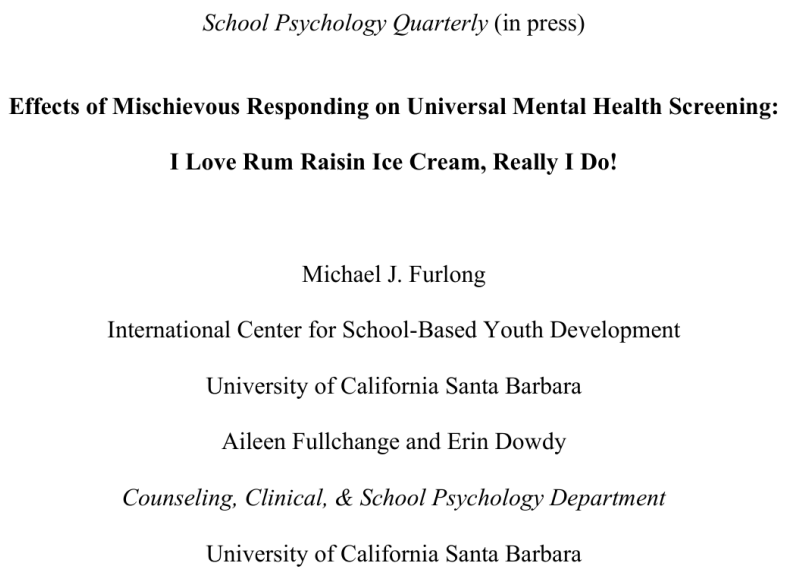
1040-3590/12/$12.00 DOI: 10.1037/a0024824

Effects of Validity Screening Items on Adolescent Survey Data

Dewey Cornell, Jennifer Klein, Tim Konold, and Francis Huang
University of Virginia

Validity screening is important!

- Can distort results
- There are several papers that discuss its importance



School Psychology Quarterly (in press)

Effects of Mischievous Responding on Universal Mental Health Screening:

I Love Rum Raisin Ice Cream, Really I Do!

Michael J. Furlong

International Center for School-Based Youth Development

University of California Santa Barbara

Aileen Fullchange and Erin Dowdy

Counseling, Clinical, & School Psychology Department

University of California Santa Barbara

ORIGINAL ARTICLES

Questionnaire and Interview Inconsistencies Exaggerated Differences Between Adopted and Non-Adopted Adolescents in a National Sample

Xitao Fan PhD, Brent C. Miller PhD, Mathew Christensen MS, Kyung-Eun Park MS, Harold D. Grotevant PhD, Manfred van Dulmen PhD, ...show all

Pages 7-27 | Published online: 23 Sep 2008

Administered using Qualtrics-- some other considerations (cont.)--

- How fast should the survey be to complete (inattentive responders; rapid guessing)?
- Middle schoolers took around 17 min, High schoolers, 12 min.
- How to determine thresholds?
- Could just be someone opening the survey to inspect and take an initial look
- Careless or inattentive responders are common with the use of Internet surveys (e.g., 10%), have been shown to reduce reliability estimates, and result in erroneous factor analytic results (J. A. Johnson, 2005; Meade & Bartholomew, 2012)
- In our sample, we only excluded ~0.7% to 2.4% of respondents (after removing invalid responders)
- Also, see Biemann et al. (2025): https://journals.sagepub.com/doi/full/10.1177/10944281251334778

The original (student) study was published in 2014

- Many more studies followed to explore different aspects of the survey

School Psychology Quarterly
2014, Vol. 29, No. 3, 238–255

1045-3830/14/$12.00 DOI: 10.1037/spq0000062

Multilevel Multi-Informant Structure of the Authoritative School Climate Survey

Timothy Konold, Dewey Cornell, Francis Huang, Patrick Meyer, Anna Lacey, Erin Nekvasil, Anna Heilbrun, and Kathan Shukla
University of Virginia

The Authoritative School Climate Survey was designed to provide schools with a brief assessment of 2 key characteristics of school climate—disciplinary structure and student support—that are hypothesized to influence 2 important school climate outcomes—student engagement and prevalence of teasing and bullying in school. The factor structure of these 4 constructs was examined with exploratory and confirmatory factor analyses in a statewide sample of 39,364 students (Grades 7 and 8) attending 423 schools. Notably, the analyses used a multilevel structural approach to model the nesting of students in schools for purposes of evaluating factor structure, demonstrating convergent and concurrent validity and gauging the structural invariance of concurrent validity coefficients across gender. These findings provide schools with a core group of school climate measures guided by authoritative discipline theory.

Keywords: school climate, multilevel factor analysis, engagement and bullying

Supplemental materials: http://dx.doi.org/10.1037/spq0000062.supp

The teacher (staff) surveys were also investigated…

JOURNAL OF SCHOOL HEALTH

American SCHOOL HEALTH Association

RESEARCH ARTICLE

Multilevel Factor Structure and Concurrent Validity of the Teacher Version of the Authoritative School Climate Survey

FRANCIS L. HUANG, PhD[a] DEWEY G. CORNELL, PhD[b] TIMOTHY KONOLD, PhD[c] JOSEPH P. MEYER, PhD[d] ANNA LACEY, MEd[e] ERIN K. NEKVASIL, MEd[f] ANNA HEILBRUN, MA[g] KATHAN D. SHUKLA, MSc, MEd[h]

Journal of Psychoeducational Assessment
1–14

DOI: 10.1177/0734282915621439
jpa.sagepub.com
SAGE

Multilevel Factor Structure, Concurrent Validity, and Test–Retest Reliability of the High School Teacher Version of the Authoritative School Climate Survey

Francis L. Huang[1] and Dewey G. Cornell[2]

Agenda

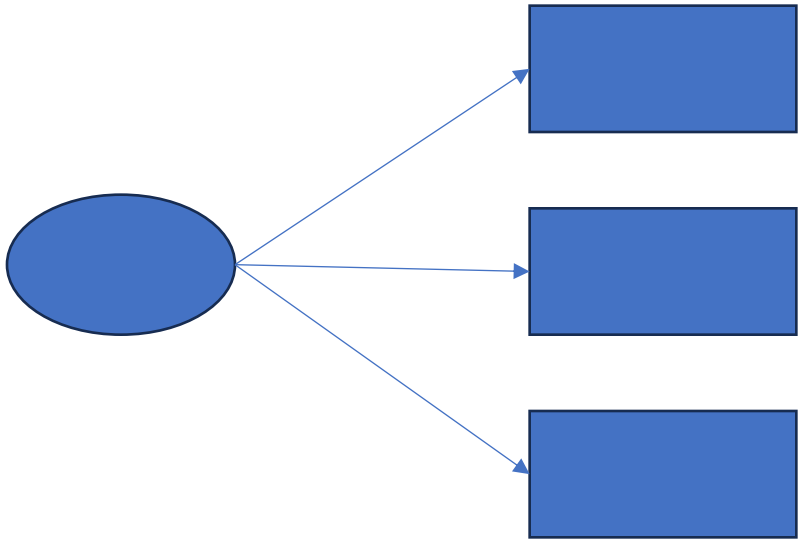
- Conceptual and theoretical background
- Survey development and administration
- Assessing psychometric properties
- Refinement and testing

Consider the measurement issues…

- School climate, by its definition, is a property of the school
- Usual tools that have been used to investigate instrument properties were meant for single-level analysis
- Need to consider the multilevel nature of how data are gathered and how items are worded (i.e., the target of the question is geared towards the school)
- Students and staff are reporters of a group-level construct– nested data
- Multilevel factor analysis was used– (back then), only Mplus could estimate this using categorical data

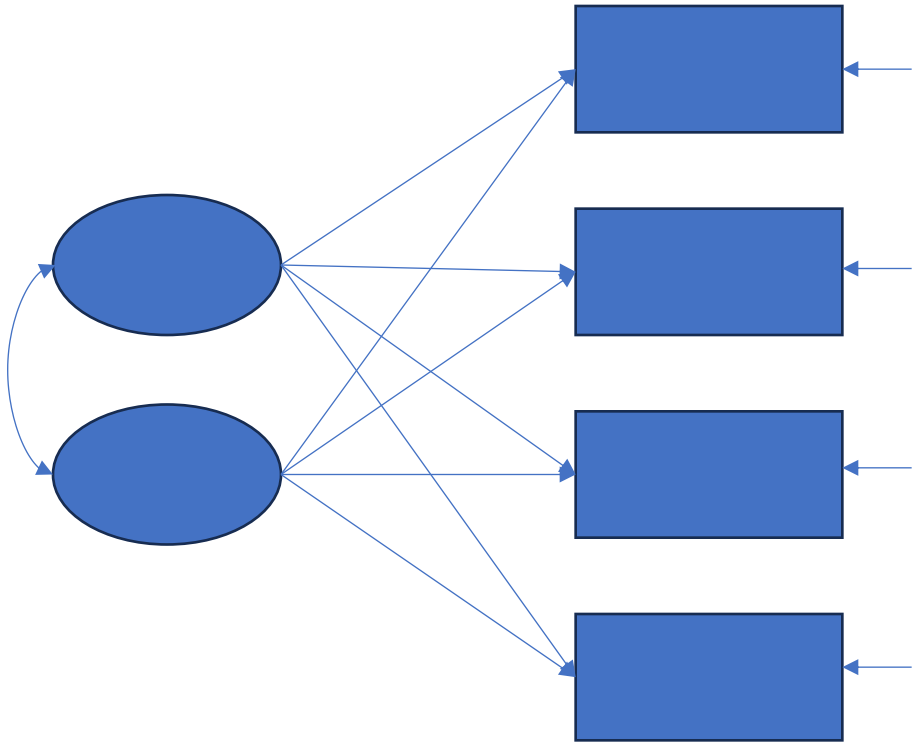
Quick refresher on factor analytic techniques…

- Idea is guided by the idea that a latent (unobserved) factor is giving an effect of how items are scored (note direction of arrows)
- Reduces data to a smaller number of dimensions or factors
- Used to identify latent constructs
- Uses the (polychoric) covariance/correlation matrix for categorical data
- Items should be correlated– the items should be correlated because of that common factor



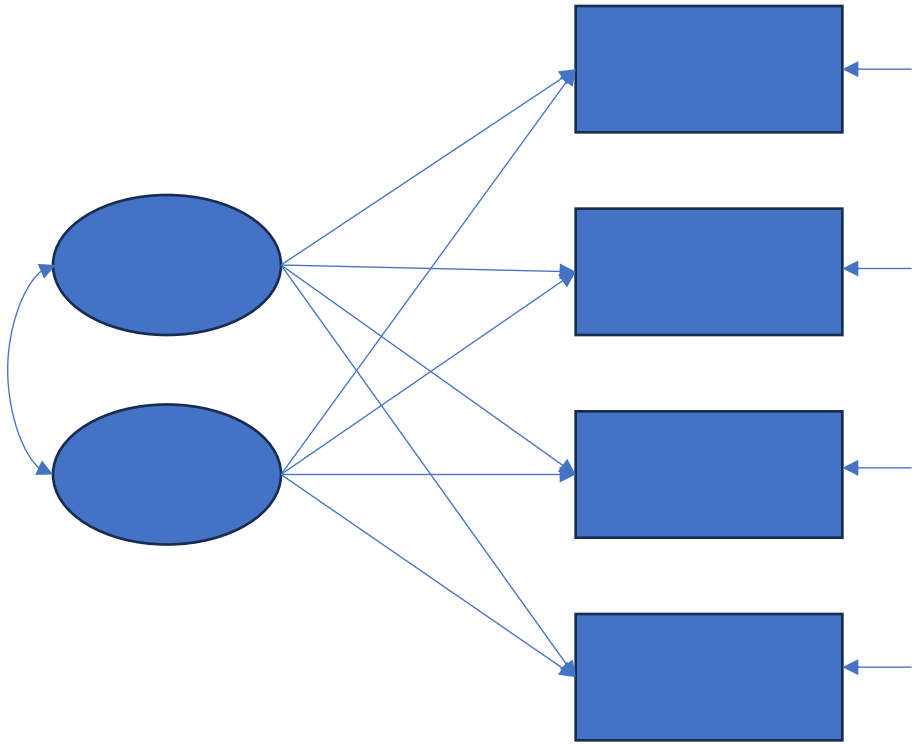
Factor structure was investigated using split-half (WHY) exploratory and confirmatory samples

- Can first start with an EFA: every item loads onto every factor
- Results are completely based on the data (without imposing any theory or constraints): just select all your variables and specify how many factors
- CFA gets at the dimensionality of the construct– it is not enough to compute the usual alpha reliabilities– does not tell you about the dimensionality!
- There were tests to determine if your data area 'suitable' for a FA (KMO, Bartlett)– but I do not bother with these– you had some idea you had a scale to begin with
- An EFA is not a PCA



Remember, with an EFA, there are many choices, such as:

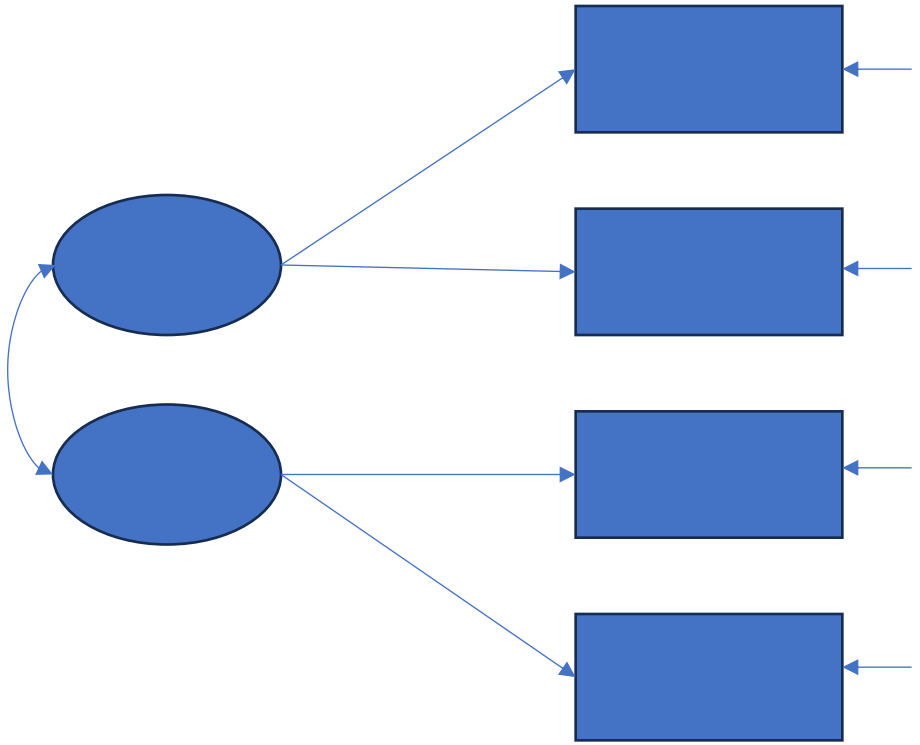
- Do you allow for factor rotation (realistically, yes)
- How many factors? (PA)
- What kind of extraction method?
- What is a meaningful loading?
- Minimum number of items per factor?



- This is a chance to revisit your items (Remove? Combine? Reword?)
- This is NO guarantee that you will get a good fit when you run a CFA

With a CFA, you impose stricter assumptions…

- What estimation method to use?
- For example, items only load onto certain factors
- Gives you a chance to revisit your factors
  - Can consult modification indices
  - Can correlate factor loadings
  - However, these warnings suggest something is going on with your factor structure
- Take into consideration fit indices (w/c can also be done with an EFA using ML)



Factor structure was investigated using split-half exploratory and confirmatory samples (cont.)

First started with an EFA

- F
- M
- N

Re
the

Fa
no
yo

Upon reflection:

- An option would have been to perform a CFA on the exploratory sample
- Consult modification indices
- Since we had theory guiding what items in the scale went together, we understood what items went together and a theory we were testing
- Then apply model to the confirmatory sample for generalizability
- I've done this for other measures afterwards…

Multilevel factor analysis (FA) is an extension of the usual FA technique but used with nested data

- Variability is decomposed in two parts:
  - Within group (level-1)
  - Between group (level-2)
- Based on separate within and between group covariance matrices– they do not have to be the same!
- Can have a different factor structure at the unit vs group level!

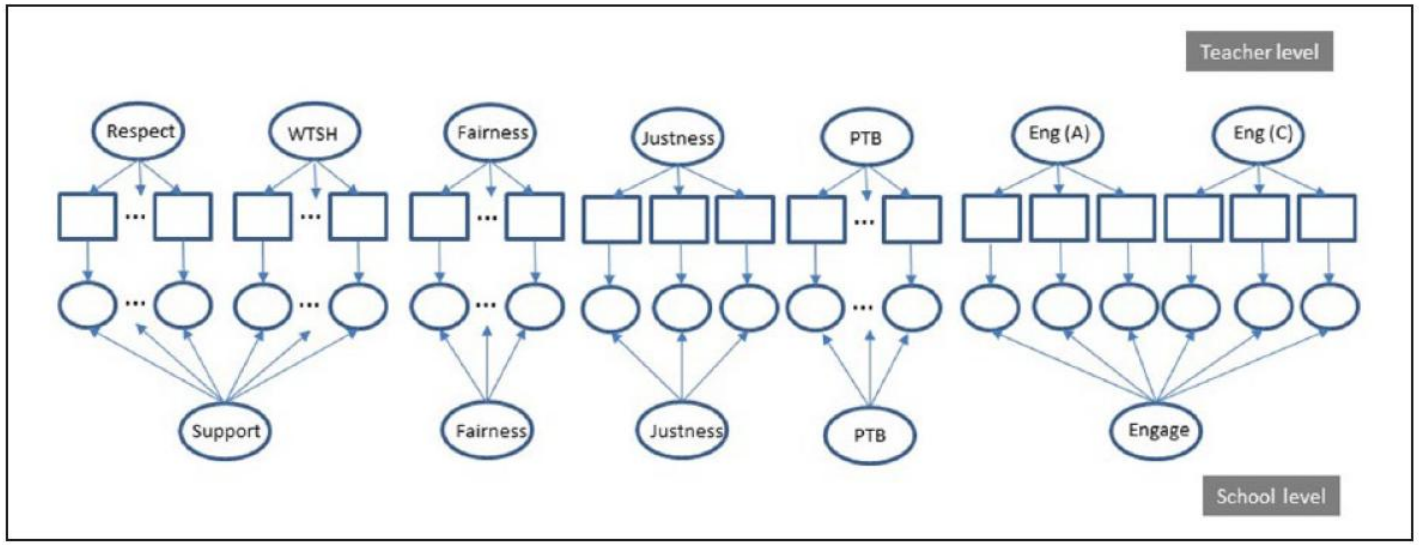


Figure 1. Conceptual Multilevel Confirmatory Factory Analysis Model.
Note. For clarity, factor correlations (refer to Tables 2 and 3) and residual variance are not shown. WTSH = willingness to seek help; PTB = prevalence of teasing and bullying; ENG (A) = affective engagement; ENG (C) = cognitive engagement.

Structure and Support were measured using several items

Disciplinary Structure

The punishment for breaking school rules is the same for all students
Students at this school only get punished when they deserve it
Students are treated fairly regardless of their race or ethnicity
Students get suspected without good reason (reverse scored)
The adults at this school are too strict (reverse scored)
The school rules are fair
When students are accused of doing something wrong, they get a chance to explain it

Student Support

Most teachers and other adults at this school care about all students
Most teachers and other adults at this school want all students to do well
Most teachers and other adults at this school listen to what students have to say
Most teachers and other adults at this school treat students with respect

There are adults at this school I could talk with if I had a personal problem
If I tell a teacher that someone is bullying me, the teacher will do something to help
I am comfortable asking my teachers for help with my school work
There is at least one teacher or another adult at this school who really wants me to do well

Factor structure was investigated using split-half exploratory and confirmatory samples (results)

Structure					Support								
	EFA	Multilevel CFA				EFA		Multilevel CFA					
	One factor	One factor				Two factors		One factor		Two factors			
						F1	F2			F1		F2	
		W	B	ICC				W	B	W	B	W	B	ICC
Structure					Support									
STR 1	.71	.68	.77	.04	SPRT 1	**.78**	.14	.86	.97	.87	.98	—	—	.06
STR 2	.60	.58	.95	.02	SPRT 2	**.66**	.28	.81	.98	.81	.99	—	—	.04
STR 3	.63	.61	.83	.05	SPRT 3	**.89**	−.02	.83	.93	.84	.95	—	—	.06
STR 4	.57	.54	.94	.06	SPRT 4	**.88**	.00	.85	.96	.86	.98	—	—	.05
STR 5	.53	.47	.83	.04	SPRT 5	.01	**.71**	.57	.69	—	—	.63	.72	.03
STR 6	.74	.72	.95	.06	SPRT 6	.31	**.49**	.66	.88	—	—	.77	.91	.05
STR 7	.63	.62	.92	.05	SPRT 7	.18	**.52**	.57	.64	—	—	.66	.67	.02
					SPRT 8	−.05	**.74**	.51	.67	—	—	.58	.70	.02

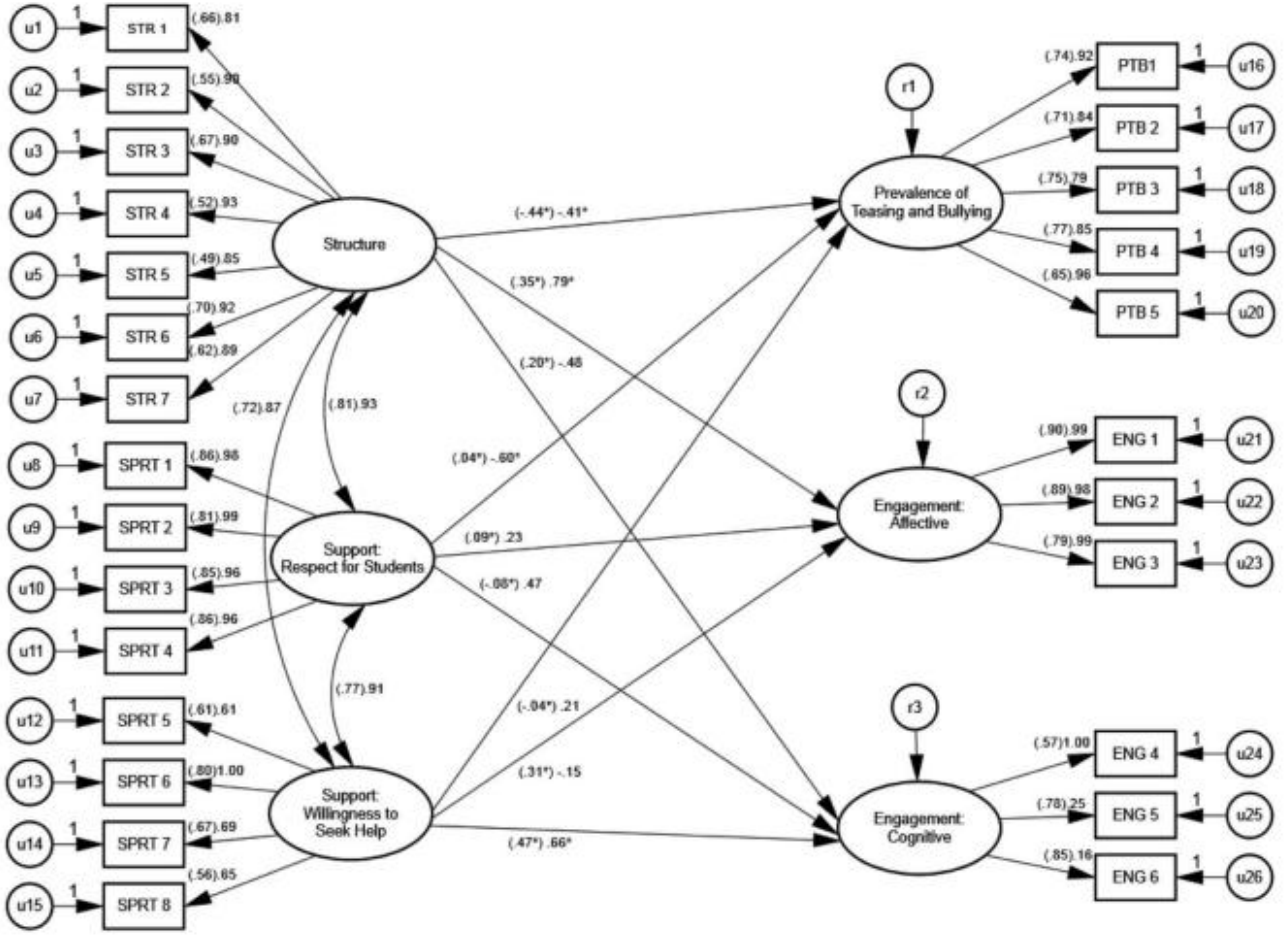
*Back then, we had used a single-level EFA– could have used a multilevel EFA. See original paper for full results.

Factor structure was investigated using split-half exploratory and confirmatory samples (additional scales)

	PTB					Engagement							
	EFA	Multilevel CFA				EFA		Multilevel CFA					
	One factor	One factor				Two factors		One factor	Two factors				
						F1	F2		F1		F2		
		W	B	ICC				W	B	W	B	W	B	ICC
PTB					Engagement									
PTB 1	.70	.73	.81	.11	ENG 1	**.91**	−.01	.89	.97	.90	.97	—	—	.11
PTB 2	.78	.69	.84	.05	ENG 2	**.93**	.00	.88	.99	.90	.99	—	—	.11
PTB 3	.80	.75	.81	.04	ENG 3	**.75**	.08	.73	1.0	.77	1.01	—	—	.06
PTB 4	.82	.77	.96	.05	ENG 4	.12	**.47**	.40	.61	—	—	.54	1.62	.05
PTB 5	.70	.69	.97	.08	ENG 5	−.01	**.98**	.62	.11	—	—	.83	.22	.03
					ENG 6	.14	**.66**	.62	.02	—	—	.83	.05	.03

*Prevalence of teasing and bullying. See original paper for full results.

SEM showing the relationship of climate measures with outcomes of bullying and engagement (zoom in…)



*Had good measures of model fit

Instrument development is an iterative process

- Scale might be better if item(s) is deleted/reworded
- Often, there is a back-and-forth between different types of model specification
- At times, researchers will correlate the residual errors but then that may just signify that an item is redundant or could be separate factors
- FA should guide construct development (do not disregard errors or warnings!)

Computing reliability is just one part of scale development

- The degree to which the instrument produces consistent results
- Often Cronbach's alpha has been misused this way
  - Alpha is **not** a test of dimensionality!
- Should consider multilevel alternatives
- Should also consider using Omega vs Alpha
  - Alpha assumes tau equivalence (same loading; not realistic)
  - With Omega, items can load differently

What about validity? Are you measuring what you think you are measuring? Many different kinds…

- Theory is your guide (as well as experts) in thinking about content validity
- Examine correlations with other measures we consider to be related (or not related)
  - Concurrent (same time)
  - Predictive (future time)

Although scales are distinct…

- May standardize and combine to avoid multicollinearity
- We know that these are separate and distinct factors

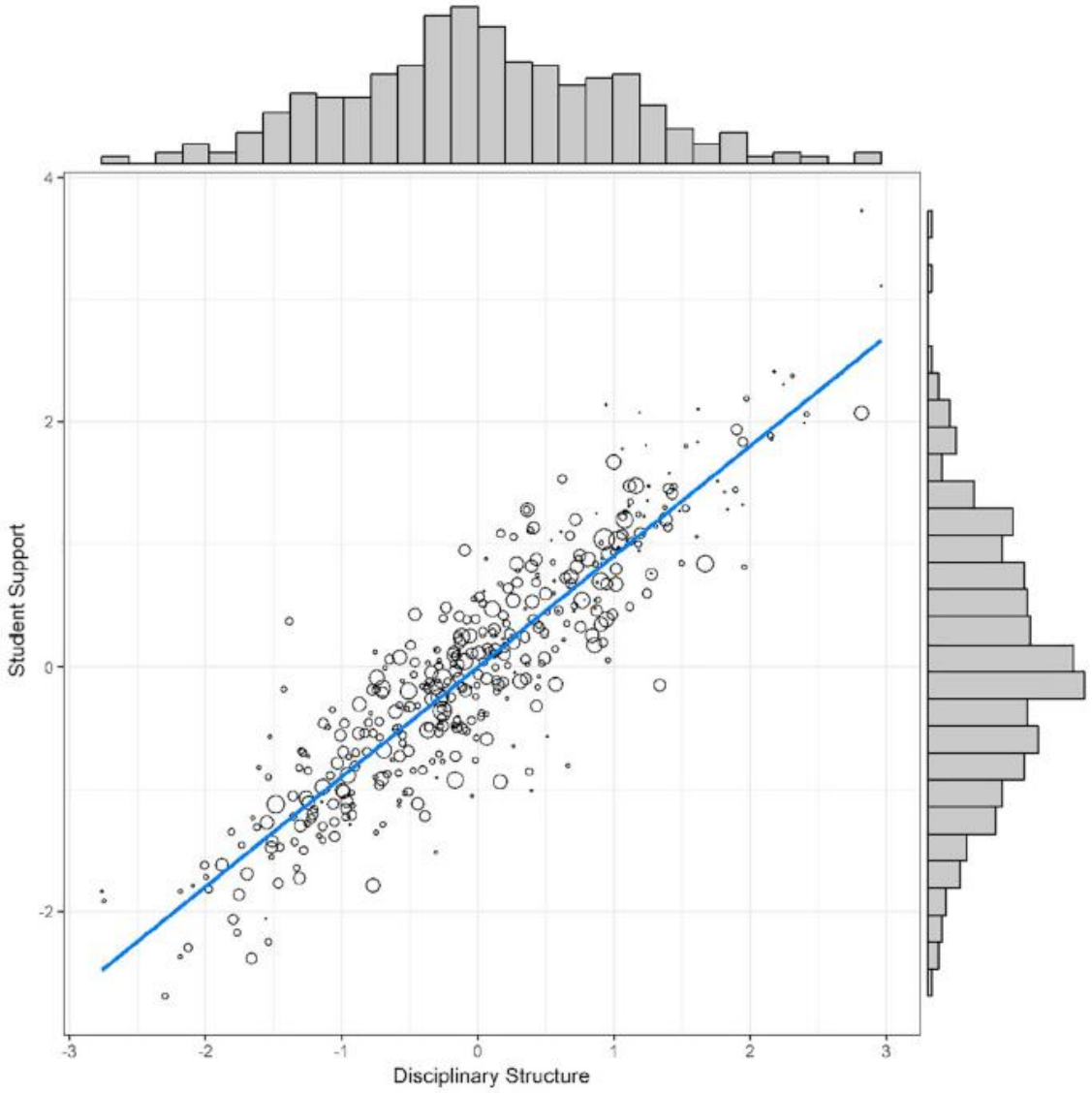


Fig. 1. Relationship between school-level disciplinary structure and student support (r = 0.88) (n = 302 schools).
Notes. Measures are standardized (M = 0, SD = 1). Size of the points reflect the school enrollment size.

Agenda

- Conceptual and theoretical background
- Survey development and administration
- Assessing psychometric properties
- Refinement and testing

Other psychometric considerations– test-retest reliability

- Test-retest reliability: if the same person took the survey again, would scores be similar?
- How much time? Two weeks
- Logistically challenging- need a way to identify responses
- Need to have a way to provide the incentive
  - In our case, it was guaranteed ($10?)
- Done for students and teachers
  - ~500 teachers were offered, 95% agreed
- Acceptable test-retest (all *r*s > .70)

Other psychometric considerations- measurement invariance

- Invariance testing: Are we measuring the construct the same way? Multiple steps…
  - By gender & race/ethnicity (multigroup invariance testing)
  - By time (temporal stability)
- Test multiple models, each with additional constraints
- When placing constraints, does the model still fit well?
- Results indicated scales were invariant

Longitudinal Measurement Invariance of the Authoritative School Climate Survey

Tim R. Konold[1], Kelly D. Edwards[1], and Dewey G. Cornell[1]

Journal of Psychoeducational Assessment
2021, Vol. 39(6) 651–664

DOI: 10.1177/07342829211011332
journals.sagepub.com/home/jpa
SAGE

There are several aspects of the survey that have been tested as well… (exploratory scales)

Aggressive Attitudes in Middle Schools: A Factor Structure and Criterion-Related Validity Study

Assessment
2015, Vol. 22(4) 497–512

DOI: 10.1177/1073191114551016
asm.sagepub.com
SAGE

Francis L. Huang[1], Dewey G. Cornell[2], and Timothy R. Konold[2]

Using Multilevel Factor Analysis With Clustered Data: Investigating the Factor Structure of the Positive Values Scale

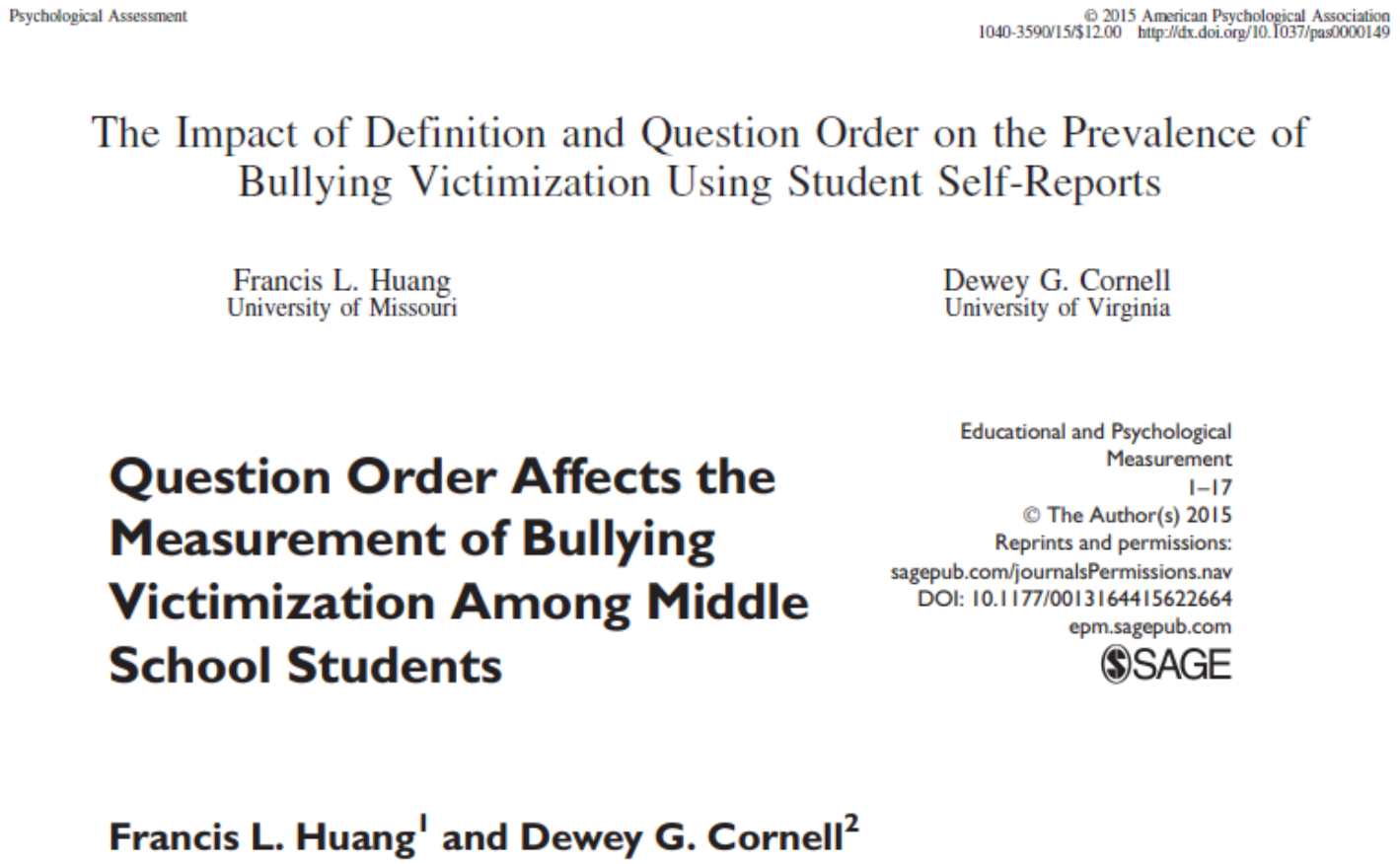
Journal of Psychoeducational Assessment
2016, Vol. 34(1) 3–14

DOI: 10.1177/0734282915570278
jpa.sagepub.com
SAGE

Francis L. Huang[1] and Dewey G. Cornell[2]

There are several aspects of the survey that have been tested as well… (cont.)

- Studies and made use of the randomization feature of Qualtrics
- In that manner, surveys are an excellent tool for experiments
  - Investigated also the use of the use of scales or sliders (anchoring effects)



Psychological Assessment

1040-3590/15/$12.00 http://dx.doi.org/10.1037/pas0000149

The Impact of Definition and Question Order on the Prevalence of Bullying Victimization Using Student Self-Reports

Francis L. Huang
University of Missouri

Dewey G. Cornell
University of Virginia

Question Order Affects the Measurement of Bullying Victimization Among Middle School Students

Educational and Psychological Measurement
1–17

DOI: 10.1177/0013164415622664
epm.sagepub.com
SAGE

Francis L. Huang[1] and Dewey G. Cornell[2]

Have used the ASCS in other states as well…

SCHOOL PSYCHOLOGY REVIEW
https://doi.org/10.1080/2372966X.2022.2109058

NASP

Check for updates

… (in Missouri and Oklahoma) …

Potential Moderation Across Racial Groups in Perceptions of Authoritative School Climate and Peer Victimization and Student Engagement

Ying-Ruey Chuang, Francis Huang, Keith Herman, and Bixi Zhang

University of Missouri

… and have investigated other areas related to school climate

SCHOOL PSYCHOLOGY REVIEW
https://doi.org/10.1080/2372966X.2023.2176160

NASP

Check for updates

The Seasonality of School Climate

Francis L. Huang, Bixi Zhang, Wendy Reinke, Keith Herman, and James Sebastian

University of Missouri

Summary

- Scale development is an iterative process
- Based on theory and prior studies
- Many considerations to think of (as shown!)
- Takes a lot of work! Very many details to consider
  - Every choice can have an effect
- Multilevel factor analytic approaches should be used when interest is in a group-level factor


Teacher Working Conditions and
their Ability to Handle Bullying

QUESTION & ANSWER

Instrument (scale) development considerations: A case study of the Authoritative School Climate Survey (ASCS)

2025.11.14

Francis L. Huang / huangf@missouri.edu
Professor / Statistics, Measurement, and Evaluation in Education Program
Educational, School, and Counseling Psychology, University of Missouri
https://francish.net
@flhuang

College of Education & Human Development
University of Missouri

Presentation @Education University of Hong Kong