Routledge
Taylor & Francis Group

Communication Education

ISSN: (Print) (Online) Journal homepage: https://nca.tandfonline.com/loi/rced20

# Instructor strictness: instrument development and validation

T. Kody Frey & Nicholas T. Tatum

**To cite this article:** T. Kody Frey & Nicholas T. Tatum (2022) Instructor strictness: instrument development and validation, Communication Education, 71:4, 327-354, DOI: 10.1080/03634523.2022.2096246

**To link to this article:** https://doi.org/10.1080/03634523.2022.2096246

COMMUNICATION EDUCATION
2022, VOL. 71, NO. 4, 327–354
https://doi.org/10.1080/03634523.2022.2096246



# Instructor strictness: instrument development and validation

T. Kody Frey [a] and Nicholas T. Tatum [b]

[a]School of Information Science, University of Kentucky, Lexington, KY, U.S.A.; [b]Department of Communication Studies, University of Nevada, Las Vegas, NV, U.S.A.

**ABSTRACT**

Three studies ($N = 1,346$) detail the development of three theoretically grounded instruments operationalizing *instructor strictness*. Using open-ended questionnaire data ($n = 427$), study 1 inductively derives an understanding of the instructor behaviors that students perceive as strict. These patterns of behavior are then condensed into a comprehensive item pool designed to measure the relevant constructs. Study 2 ($n = 391$) evaluates the underlying factor structures comprised by the patterns of strictness identified in study 1 through a series of exploratory factor analyses. Study 3 ($n = 528$) establishes factorial validity of each new measure through confirmatory factor analyses. Studies 2 and 3 also provide evidence for convergent and concurrent validity between the newly formed measures of *evaluative*, *regulatory*, and *interactive* strictness and relevant variables within the nomological network, including the Questionnaire on Teacher Interaction (QTI) strictness inventory, the QTI admonishing behavior inventory, instructor caring, and cognitive flexibility. The research provides a roadmap to investigate how instructors who enforce classroom rules or demonstrate inflexibility may influence instructional outcomes in nuanced ways. The theoretical and practical implications of the new measures for instructional communication research, as well as future directions, are discussed.

**ARTICLE HISTORY**
Received 15 March 2022
Accepted 27 June 2022

**KEYWORDS**
strictness; inflexibility; policies; classroom management

Students routinely use the word "strict" to describe instructor behavior in the classroom (Jenkins, 1997; Poplin et al., 2011; Wubbels et al., 1991), and variations of the term appear frequently throughout popular culture (e.g., Carr, 2014; Tickle, 2017). In terms of scholarship, strictness is routinely discussed across disciplines (e.g., Jiang et al., 2021; Wubbels & Levy, 1991); however, no apparent research considers instructor strictness as a variable of interest or systematically explores the role it could play in the classroom (Tatum & Frey, 2021). Instead, because of its colloquial pervasiveness, the term strictness is often used without consideration for how it is actually managed and prescribed in classroom contexts. For example, in their study of instructor self-disclosure online, Mazer et al. (2007) claimed that " … teachers who exhibit a relaxed personality on Facebook with informal photographs and entertaining messages, but show themselves to be strict in the classroom, may create violated student expectations that may result in negative effects on students" (p. 13). Likewise, Larseingue et al. (2012) argued that "teachers

**CONTACT** T. Kody Frey tkfr222@uky.edu

can modulate the perceived rigor of communication courses by regulating the amount of work they require as well as by adjusting the strictness with which they adhere to grading standards" (p. 438). Such examples may be intuitively understood, yet the true impact of strict instruction remains anecdotal until scholars solidify and empirically validate this often expressed, but seldom clarified, concept.

The need for research exploring strictness becomes even more apparent when considering the impact of the 2020 SARS-CoV-2 (COVID-19) pandemic. The virus uprooted standards for instructor behavior in a variety of ways and raised ethical questions regarding how pedagogy and instruction should be enacted (Darling-Hammond & Hyler, 2020). The conversation surrounding how to put students in the best position to learn despite increasing worries about health care, finances, and personal safety, among other topics, toed the line between offering increased accommodations and remaining firm in the face of uncertainty. As Tatum and Frey (2021) argued, "the pandemic has led to a clamoring of voices making broad, sweeping generalizations concerning what students need and want in terms of strictness" (p. 215). However, the efficacy of such claims needs to be called into question without empirical investigations that explore the effects of this flexible, or inflexible, behavior.

The present studies seek to address these gaps, elaborating on the conceptualization of instructor strictness and developing three reliable and empirically derived instruments to assess instructor strictness in the college classroom. Study 1 employs open-ended survey data to inductively develop a typology of instructor messages and behaviors related to perceived strictness as the basis for scale item generation. Study 2 submits the items generated from study 1 to exploratory factor analysis (EFA) procedures to reduce them to a meaningful set of underlying factors. Additionally, this study examines convergent validity between the newly formulated measures and existing inventories of strictness and admonishing behavior. Finally, Study 3 submits each measure to confirmatory factor analysis (CFA) procedures to holistically test the data against a previously theorized factor structure. Study 3 also provides tests of concurrent validity through the expected relationship between instructor strictness and other theorized constructs (i.e., instructor caring, cognitive flexibility).

## Conceptualizing instructor strictness

The first step in truly understanding a construct should be the formation of a precise definition within the desired context (Hinkin, 1998). Drawing from several descriptions

**Table 1.** Existing descriptions of strict instructor behavior.

| Article | Definition |
|---|---|
| Coll et al. (2002) | "Severe when marking papers, standards are very high, tests are hard, afraid of this teacher, silent in this teacher's class" (p. 120) |
| Den Brok and Brekelmans (2001) | "Keeps a tight rein, checks, judges, demands silence, sets rules, gives hard tests" (p. 25) |
| Fisher et al. (1995) | "The extent to which teachers keep the reins tight, maintain silence and consistently check the rules" (p. 132) |
| Kendall and Schussler (2013a) | "Adhere to policies and rules, are inflexible, are tough graders, and not tolerant of bad behavior or distractions" (p. 205) |
| Wubbels et al. (1991) | "Keep reins tight, check, judge, get class silent, maintain silence, be strict, exact norms and set rules" (p. 3) |

of strict instructors (see Table 1), Tatum and Frey (2021) offered the following definition of strictness: "the perceived inflexibility of an instructor based on their unwavering adherence to instructional policies and procedures" (p. 2). Herein, this definition is expanded, and its assumptions are elucidated to create a more nuanced understanding of the construct of interest.

*First*, strictness is enacted by instructors through the enforcement of implicit and explicit policies and procedures in the classroom. Broadly speaking, policies and procedures refer to an instructor's expectations for how classrooms should be managed in terms of formal classroom rules (e.g., the syllabus; Thompson, 2007) or enacted classroom norms (e.g., gives hard tests; Den Brok & Brekelmans, 2001). Regardless of format (e.g., face-to-face, hybrid, online), classroom policies typically exist to prepare instructors "for what occurs when a class begins and so they can decrease the likelihood that students will feel the need to engage in aversive behaviors" (Hayes, 2005, p. 62). The nature and importance of these policies and procedures are unique to each classroom. In many cases, instructors have autonomy to manage their classrooms how they see fit. However, various programs, departments, and universities might enforce universal policies (e.g., attendance, Title IX). So, the first assumption within this definition is that strictness occurs in reference to a wide range of rules and norms, many of which may not be applicable to every instructional setting or may be prescribed by entities other than an instructor.

*Second*, for an instructor's behavior to be described as strict, it must be perceived as "inflexible" by students. Thus, not all adherence to a policy or a procedure is considered strict—only that which is perceived to be keeping reins tight. In this way, one student may perceive adherence to be lenient while another could view the same behavior as rigid. Flexibility as a communicative construct has been explored in the discipline at large (Martin & Rubin, 1994) and is often considered "an essential component of communication competence" (Martin & Anderson, 1998, p. 2). More specifically in the classroom, researchers have identified flexibility, or lack thereof, as a key communication behavior that drives students' perceptions such as instructor credibility (Myers & Bryant, 2004) and teaching evaluations (Worley et al., 2007).

In fact, in a classroom setting, students identify instructors who demonstrate inflexibility through unreasonable and arbitrary rules as misbehaving. In the initial investigation of instructor misbehaviors (Kearney et al., 1991), students specifically described instructors who refused to accept late work, gave no breaks during long classes, punished entire classes for a single student's behavior, or generally behaved with rigidity, inflexibility, or authoritarianism as interfering with learning. Subsequent replications in more current classroom contexts (Goodboy & Myers, 2015), as well as online learning environments (Vallade & Kaufmann, 2018), have also identified inflexibility stemming from unreasonable and arbitrary rules as misbehaviors. Collectively, the second assumption within this definition is that strictness does not explicitly refer to the characteristics or features of policies and procedures but instead reflects the extent to which the instructor adheres to them. This means that simply having a specific policy does not constitute strictness; for a behavior to be strict, it must involve unwavering adherence to a policy (or contrarily, flexible adjustment away from stated or implied expectations).

*Third*, strictness does not necessarily produce unfavorable evaluations, as the definition forwarded by Tatum and Frey (2021) is void of either positive or negative

valence. Despite the notion that a theoretical grounding in flexibility is likely to lead students to perceive inflexible instructors undesirably, there may be some conditions whereby it has the opposite effect. Indeed, a strict instructor "may be unbending in their rules to the point of being unfair" (Kendall & Schussler, 2013b, p. 102), and this unfairness will likely be viewed negatively (see Chory-Assad & Paulsel, 2004). But it is also entirely possible that strict instructors may be perceived favorably when enforcing "polite behavior in a large classroom" or even applying "the same rules to all students" (Kendall & Schussler, 2013b, p. 102). Some instructors have indeed experienced success through adherence to policies or procedures they view as strict (for example, see Trefzger, 2018). Moore and Richards (2019) extended this line of thinking by suggesting that "favorable student outcomes may occur in response to reward-based, prosocial strategies *and* to punishment-based, antisocial strategies, depending upon the credibility of the instructor and the framing of the syllabus policy" (p. 405). Thus, on its own, strictness is likely to be perceived negatively; however, depending on how an instructor frames their inflexibility or applies their rules, some students could benefit from, or even appreciate, strict instruction. The third assumption within this definition is that instructor strictness is not value-laden; it describes the strict messages and behaviors themselves and excludes students' evaluations of them.

## Operationalizing instructor strictness

Based on this expanded conceptualization of strictness, no existing instruments appear to reflect the full scope of the construct in the twenty-first-century classroom. In the few times it has been measured, the concept has only been captured using a small number of items from a larger scale. Most predominantly, six items from Wubbels et al.'s (1985) 48-item Questionnaire on Teacher Interaction (QTI) have been employed to elicit students' perceptions of an instructor's strictness: (1) this teacher is severe when marking papers; (2) this teacher's standards are very high; (3) we have to be silent in this teacher's class; (4) we are afraid of this teacher; (5) this teachers' tests are hard, and (6) this teacher is strict. For several reasons, these items are inadequate for operationalizing strictness. The items do not reference important, mandated policies and procedures found in most classrooms like those related to academic integrity or deadlines. Further, the items make no mention of technology restrictions, which are pervasive in modern classrooms (e.g., Tatum et al., 2018). Finally, it seems improbable that such a limited number of items, especially given that the QTI was not developed following modern instrument development standards, truly capture the scope of this abstract concept. As such, there is an exigency to develop a novel instrument to measure instructor strictness so it can be explored quantitatively in research.

## Study 1: thematic analysis and instrument development

Given the arguments presented for the conceptualization of strictness, study 1 sought to understand students' actual, lived experiences with strict instruction. Furthermore, analyzing students' actual recollections of their interactions with strict instructors should overcome the limitations posed by existing scales by revealing specific instances of

inflexible behavior that better reflect the current educational context. In turn, these instances can be used as the basis for item generation in the construction of a measure. The following research question was posed:

> RQ1: Given the present definition, what types of instructor messages or behaviors do students perceive as strict?

## Participants and procedures

Participants included 427 undergraduate students enrolled at a large institution in the Southeast (160 men, 37.47%; 260 women, 60.89%, seven did not report, 1.64%). Participants ranged in age from 18 to 54 ($M = 19.69$, $SD = 2.87$), with 174 first-year students, 57 sophomores, 108 juniors, 81 seniors, one fifth-year senior or beyond, and six who did not report. Students identified as White/Caucasian ($n = 329$; 77.05%), Black or African American ($n = 30$; 7.03%), Asian or Pacific Islander ($n = 36$; 8.43%), Hispanic or Latino/Latina ($n = 16$; 3.75%), Native American or American Indian ($n = 1$; 0.23%), and Other ($n = 9$; 2.11%), with 6 (1.41%) not reporting.

Following approval from the university IRB, students were recruited from the introductory communication course via a research participation system. Participants who self-selected to participate were provided with a link to a survey hosted by Qualtrics. After agreeing to participate, students were then directed to respond to several open-ended and demographic questions. Students who completed the survey earned minimal course credit (1% of their total grade). Those choosing not to take part in the research were provided with an alternate assignment, providing them with the opportunity to earn the same amount of credit. Students were first provided with the definition of instructor strictness forwarded by Tatum and Frey (2021). Then, they were prompted to provide written narratives about their experiences with instructor strictness in the classroom. Specifically, participants responded to two open-ended questions designed to elicit a wide variety of experiences and perspectives: (1) *Tell me a story about when an instructor during college was strict*; and (2) *What other things do you believe an instructor can do to make them seem strict*? Students were not asked to report on a specific instructor, and no data was collected regarding the instructors students were referencing in their open-ended responses.

## Thematic analysis

Following analytic procedures in similarly derived scale development studies (e.g., Ledbetter, 2009; Mazer, 2012), data were subjected to qualitative thematic analysis and interpretation using techniques established by Braun and Clarke (2006). Braun and Clarke defined a thematic analysis as "a method for identifying, analysing, and reporting patterns (themes) within data" (p. 79). Their approach identified six steps for a thorough analysis: (1) familiarizing yourself with the data, (2) generating initial codes, (3) looking for themes, (4) evaluating themes, (5) defining and naming themes, and (6) producing the report.

First, the authors read through the responses individually to get a better, holistic sense of the data. The authors did not rely on any pre-existing categorizations, establishing a purely etic analysis that reflects Braun and Clarke's (2006) constructivist perspective. Second, after the individual read-through, the authors met to compare their initial interpretations of the

data. This involved identifying similarities and differences among the responses and identifying individual units. Third, the authors identified key themes and concepts from the data. The researchers read the data line by line to establish categories in the form of theoretical concepts, events, or markers based on their surface-level characteristics (Lincoln & Guba, 1985). Fourth, authors met on several occasions to modify the existing categories based on their understanding of the data or add new categories for responses that did not fit an existing category. Fifth, authors combined the concepts from the surface-level categories to reflect themes occurring broadly across the data set (Lindlof & Taylor, 2017). Authors utilized the constant comparative method until they reached consensus that the categories and themes accurately described the data. The analysis resulted in the identification and naming of three unique themes reflecting separate types of strict behavior (i.e., evaluative, regulatory, and interactive strictness). Sixth, these themes were used to generate an initial item pool for subsequent measurement (DeVellis, 2003).

## Study 1 Results

Overall, students described three patterns of strict behaviors and messages. *Evaluative strictness* refers to the enforcement of rules, policies, or practices dealing with the direct assessment of students' content knowledge or ability to complete instructional objectives. Evaluative strictness was enacted when instructors did not provide students with opportunities to improve grades, graded harshly, held high expectations for how much content students should know, assigned a heavy workload in a short timeframe, or lacked clarity when providing information necessary to complete or perform tasks. *Regulatory strictness* refers to the way instructors used policies and procedures to manage the classroom and facilitate ideal conditions for student learning and growth. Regulatory strictness occurred when instructors required attendance, refused to accept late work, banned technologies (e.g., phones, laptops), mandated specific conditions for taking tests or exams (e.g., everyone has to take the exam at the same time), enforced implicit norms for classroom behavior not stated within the syllabus (e.g., no using the bathroom, no eating in class), or did not deviate from their chosen teaching methods (e.g., they would only lecture). *Interactive strictness* refers to the way instructors interpersonally communicated about their policies and procedures to students both within and outside the classroom. Students perceived interactive strictness when instructors refused to acknowledge their reasoning for breaking a policy, publicly addressed students for not adhering to policies, or made themselves unavailable for help.

Next, an initial pool of 52 items was created from previous descriptions of strictness and the qualitative findings (Carpenter, 2018). Importantly, because the results of study 1 appeared to suggest that each pattern of strictness resembled its own subset of behaviors, a decision was made to treat the patterns of strictness as separate, independent constructs (cf., classroom justice, Chory-Assad & Paulsel, 2004). Given the unique characteristics of each pattern, the authors felt that aggregating all the items into a unified instrument may have masked interesting latent constructs or jeopardized potentially nuanced relationships with relevant classroom outcomes. Thus, 15 items were written to measure evaluative strictness, 22 items were written to measure regulatory strictness, and 15 items were written to measure interactive strictness.

Finally, a focus group of seven undergraduate students was utilized to enhance the face validity of the item pool. After discussing their personal experiences with strict instructors, students reviewed the proposed definition and list of items and completed a full version of each instrument. Following this process, one of the authors reviewed the items individually with the focus-group participants to assess their comprehensiveness and clarity. Participants offered a small number of changes regarding item phrasing, wording, and specificity. In general, the focus group determined that the items represented their understanding of the construct. For a full list of items, see Table 2.

Based on the conditions outlined in the review of literature, the thematic analysis, and a student focus group, study 1 presents three initial item pools—evaluative strictness, regulatory strictness, and interactive strictness—that align with the definition proposed by Tatum and Frey (2021). Study 2 individually tests the factor structure of the newly derived instruments and examines relationships with other operationalizations of strict instruction as a method of ensuring the proposed items are measuring the intended construct.

## Study 2: EFA and convergent validity

Study 2 refines the instruments by subjecting each set of items to EFA procedures. The study also includes an examination of relationships between each pattern of strictness and related constructs as a test of convergent validity.

### *Convergent validity*

Convergent validity is established when an instrument correlates positively with scales that measure the same or similar constructs and have been previously validated. Therefore, two scales that assess strictness or related constructs were compared to the newly derived strictness measures. First, the most frequently cited measure of instructor strictness comes from Wubbels et al.'s (1985) QTI. Second, a related aspect to strictness within the QTI is the perception of an instructor's admonishing behavior. This scale reflects the extent to which an instructor conveys anger or punishes students. Given the current conceptualization of strictness, in addition to the results of study 1 which indicate that instructors who become angry, show rudeness, express irritation, or punish students are indeed strict, this relationship was also examined. The new measures expand upon the inventories proposed by the QTI, indicating they should be related. Accordingly, the following research question and hypotheses were posed to guide study 2:

RQ2: What factor structure and reliability are present in the items for perceived (a) evaluative, (b) regulatory, and (c) interactive strictness?

H1: Each proposed measure will be positively associated with the strictness dimension of the QTI.

H2: Each proposed measure will be positively associated with the admonishing dimension of the QTI.

**Table 2.** Study 1 initial item results.

*Evaluative Strictness*
1. Refused to curve grades on assignments (e.g., projects, exams).
2. Refused to curve final grades.
3. Denied students' requests to round their grades up.
4. Allowed students to redo assignments. [R]
5. Gave severe penalties for small mistakes when grading.
6. Consistently docked students a small number of points when grading.
7. Made assignments challenging.
8. Removed points for not adhering to minor assignment details.
9. Provided adequate time to complete assignments after they were assigned. [R]
10. Required students to complete multiple assignments each week.
11. Overwhelmed students with busy work.
12. Required students to complete unnecessary assignments.
13. Provided unclear instructions for completing assignments or tasks.
14. Contradicted their own instructions or guidelines.
15. Gave assignment directions that left students confused.

*Regulatory Strictness*
1. Penalized students for being late to class.
2. Required students to attend class meetings.
3. Got angry when students came to class late.
4. Needed documentation to excuse student absences.
5. Declined to extend assignment deadlines.
6. Refused to accept late work.
7. Would not grade an assignment if it was turned in late.
8. Did not allow students to use personal devices (e.g., phones, laptops) in class.
9. Publicly called out students for using personal devices (e.g., phones, laptops) in class.
10. Punished students for inappropriately using of technology in class.
11. Controlled how students used technology (e.g., phones, laptops; cameras) in class.
12. Made students take tests and exams at the same time to prevent cheating.
13. Used technology (e.g., software, cameras) to closely monitor individual behavior when taking tests.
14. Took steps to ensure students could not cheat on tests or exams.
15. Gave students assigned seats.
16. Prohibited students from talking to one another during class.
17. Banned students from eating during class.
18. Forbade students from using the restroom during class.
19. Enforced classroom rules that were not directly stated in the syllabus.
20. Refused to deviate from the course schedule.
21. Restricted how students took notes during class.
22. Did not allow students to talk informally with one another during class.

*Interactive Strictness*
1. Refused to acknowledge a student's reasoning for breaking a policy.
2. Did not listen to student excuses.
3. Was not willing to listen to students' excuses.
4. Was understanding of a variety of student circumstances. [R]
5. Belittled students if they did not follow classroom rules.
6. Made disparaging comments toward students who did not follow classroom rules.
7. Spoke to students with a stern tone.
8. Took frustration out on students when they did not follow a rule or policy.
9. Was mean to students who did not follow rules or policies.
10. Was not available for additional help outside of set class time.
11. Showed no interest in helping students.
12. Was unavailable to help students solve problems.
13. Was unavailable to help students complete tasks.
14. Went out of their way to make themselves available to students. [R]
15. Did not desire to have a relationship with the students in class.

Note. [R] indicates that item is reverse-coded.

## *Participants and procedures*

Participants ($N = 391$) included undergraduate students enrolled in communication courses at a large, Southeastern university (148 men, 37.85%; 241 women, 61.64%; one

identified as a transgender female; 0.26%, one did not report; 0.26%). The participants' ages ranged from 18 to 55 ($M = 19.61$, $SD = 2.66$), including 158 first-year students (40.41%), 89 sophomores (22.76%), 85 juniors (21.74%), 47 seniors (12.02%), six fifth-year seniors or beyond (1.53%), and six unreported (1.53%). They reported their ethnicity as White/Caucasian ($n = 310$; 79.28%), Black or African American ($n = 34$; 8.70%), Asian or Pacific Islander ($n = 13$; 3.32%), Hispanic or Latino/Latina ($n = 16$; 4.09%), Native American or American Indian ($n = 4$; 1.02%), Bi-racial or Mixed ($n = 9$; 2.30%), Other ($n = 4$; 1.02%), and unreported ($n = 1$; 0.26%).

After receiving approval from the university's IRB, students were solicited through a research participation system in undergraduate communication courses. Students were provided with a brief description of the research which included an overview of the purpose and the amount of time they could expect for participation. Additionally, they received course credit for their involvement. Like study 1, those choosing not to take part in the study were provided with an alternative assignment for equivalent credit. Students responded to a survey administered through a secure and unique link hosted by Qualtrics. Finally, the design avoided students reporting on the same instructor by asking them to answer questions about the instructor of the course they attended prior to taking part in the research (Plax et al., 1986).

### Measures

#### Study 1 initial item results

Students responded to the 15 items assessing evaluative strictness, the 22 items assessing regulatory strictness, and the 15 items assessing interactive strictness developed in study 1. Items were created in reference to the stem "My instructor … ", and responses were drafted using a Likert-type scale ranging from *Strongly Disagree* (1) to *Strongly Agree* (7).

#### Strict behavior

The strictness dimension of the QTI uses 6 items to represent students' perceptions of the perceived strictness of an instructor. Responses were collected using a Likert-type scale ranging from 0 (*Never*) to 4 (*Always*). McDonald's omega reliability for the measure was estimated at $\omega = .841$ ($M = 2.43$, $SD = 0.84$).

#### Admonishing behavior

The admonishing dimension of the QTI uses six items to represent students' perceptions of an instructor's ability to warn or reprimand them ("This instructor gets angry too quickly"; "This instructor is quick to correct us when we break a rule"). Responses were collected using a Likert-type scale ranging from 0 (*Never*) to 4 (*Always*). McDonald's omega reliability for the measure was estimated at $\omega = .919$ ($M = 1.45$, $SD = 0.73$).

## Results

### Exploratory factor analyses

RQ2 sought to analyze the factor structure and reliability for the evaluative, regulatory, and interactive strictness instruments. To answer this question, items comprised by each

**Table 3.** EFA initial extraction model fit indices.

| Factors | Parameters | $\chi^2$ | *df* | RMSEA [90% CI] | SRMR | CFI | TLI |
|---|---|---|---|---|---|---|---|
| *Evaluative Strictness* | | | | | | | |
| 4 | 78 | 144.18 | 51 | .068 [.055, .082] | .032 | .965 | .928 |
| 3 | 67 | 248.25 | 63 | .087 [.076, .098] | .041 | .931 | .885 |
| 2 | 55 | 486.00 | 76 | .117 [.108, .128] | .065 | .848 | .789 |
| 1 | 42 | 900.22 | 90 | .152 [.143, .161] | .107 | .700 | .649 |
| *Regulatory Strictness* | | | | | | | |
| 6 | 161 | 262.62 | 114 | .058 [.049, .067] | .024 | .963 | .925 |
| 5 | 144 | 351.92 | 131 | .066 [.058, .074] | .031 | .945 | .903 |
| 4 | 126 | 491.16 | 149 | .077 [.069, .084] | .046 | .915 | .868 |
| 3 | 107 | 833.83 | 168 | .101 [.094, .108] | .073 | .835 | .772 |
| *Interactive Strictness* | | | | | | | |
| 3 | 67 | 299.41 | 63 | .098 [.087. .109] | .038 | .954 | .922 |
| 2 | 55 | 704.55 | 76 | .145 [.136, .156] | .067 | .877 | .829 |
| 1 | 42 | 1151.88 | 90 | .174 [.165, .183] | .077 | .792 | .756 |

Note. All $\chi^2$ correlations are significant at $p < .001$.

scale were subjected to EFA procedures with maximum likelihood estimation via the *psych* package (Revelle, 2017) in the free statistical software RStudio (Version 1.4.1717; RStudio Team, 2021). RStudio was chosen to conduct the analyses to allow for comparison of a variety of models using relevant fit indices; all codes are provided in the online appendix.[1] For each instrument, we followed Carpenter's (2018) advice and examined solutions ranging across several factors—as dictated by both the scree plot and parallel analysis (PA; Horn, 1965)—to determine the model that best fit the data. All models were conducted using direct oblimin rotation to allow factors to correlate. Following the guidance of Sloat et al. (2017), models were then evaluated based on a blend of (1) chi-square, (2) fit indices, (3) eigenvalues, (4) factor loadings, and (5) factor interpretability. After an appropriate model was selected, the criteria for item and factor retention were: (1) primary factor loadings of .50 or greater, (2) no secondary factor loadings exceeding .30, (3) loading on a factor with a minimum of two items, and (4) theoretical interpretability. Items that did not adequately load onto a factor were iteratively eliminated. Fit statistics across all initial extraction models are presented in Table 3.

### *Evaluative strictness*

During the initial screening, both the scree plot and the PA suggested that four factors best fit the 15 items for evaluative strictness. Thus, we examined models with solutions ranging from one to four factors. Both the KMO measure (0.88) and Bartlett's test [$\chi^2 = 2804.872$ (105), $p < .001$] were acceptable. As shown in Table 3, the four-factor model provided the best fit compared to the other models. Thus, we determined that the four-factor solution best represented the data.

After iteratively removing three items not meeting the .50/.30 criteria, the final, four-factor solution collectively explained 63.77% of the variance. The first factor was labeled *Standards* ($M = 3.40$, $SD = 1.41$, $\omega = .793$). This dimension consisted of three items relating to the amount of work students were expected to complete for the class. The second factor was labeled *Rounding* ($M = 3.83$, $SD = 1.30$, $\omega = .843$). This dimension included three items related to the instructor's willingness to curve project, exam, and final grades. The third factor was labeled *Harshness* ($M = 3.27$, $SD = 1.44$, $\omega = .827$). This dimension included three items related to the meticulousness of the instructor's

**Table 4.** Rotated factor structure of evaluative strictness.

| Survey item | Factor | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| *Standards* | | | | |
| 1. Overwhelmed students with busy work. | **.87** | .00 | .03 | −.01 |
| 2. Required students to complete unnecessary assignments. | **.77** | .01 | −.02 | .15 |
| 3. Required students to complete multiple assignments each week. | **.52** | .03 | .00 | −.07 |
| *Rounding* | | | | |
| 4. Refused to curve final grades. | −.01 | **.98** | −.07 | .03 |
| 5. Refused to curve grades on assignments (e.g., projects, exams). | −.01 | **.72** | .07 | .02 |
| 6. Denied students' requests to round their grades up. | .09 | **.58** | .25 | −.14 |
| *Harshness* | | | | |
| 7. Gave severe penalties for small mistakes when grading. | −.10 | .01 | **.84** | .13 |
| 8. Consistently docked students a small number of points when grading. | .15 | .01 | **.76** | −.07 |
| 9. Removed points for not adhering to minor assignment details. | .09 | .06 | **.57** | .05 |
| *Ambiguity* | | | | |
| 10. Provided unclear instructions for completing assignments or tasks. | .03 | .03 | −.02 | **.86** |
| 11. Gave assignment directions that left students confused. | .06 | −.01 | .15 | **.62** |
| 12. Contradicted their own instructions or guidelines. | .22 | .05 | .17 | **.51** |
| SS loading | 1.98 | 1.93 | 2.00 | 1.73 |
| Variance accounted for | 16.53% | 16.10% | 16.70% | 14.44% |

Note. SS loading = rotated sum of squares loading. Bold factor coefficients show acceptable factor loadings for the corresponding items and factors. Items that are not in bold did not load on the corresponding factor.

grading criteria. The fourth factor was labeled *Ambiguity* ($M = 2.82$, $SD = 1.46$, $\omega = .847$). This dimension included three items related to the way the instructor provided students with instructions or directions. The comprehensive scale demonstrated high internal reliability of $\omega = .881$. Using cutoff criteria suggested by Hu and Bentler (1999), the final four-factor solution fit the data well: $\chi^2(24) = 49.990$, $p = .001$; RMSEA = .053 (.032, .073); CFI = .989, TLI = .970, SRMR = .018. Factors loadings for individual items in the final model are presented in Table 4.

### *Regulatory strictness*

The scree plot suggested that five factors best fit the 22 items for regulatory strictness. Contrarily, the PA suggested a six-factor model. Thus, we examined models with solutions ranging from three to six factors. Both the KMO measure (0.89) and Bartlett's test [$\chi^2 = 4267.68$ (231), $p < .001$] were acceptable. The results demonstrated worsening of model fit when moving from the six-factor model to the five-, four-, and three-factor models. However, within the six-factor model, the two items comprised by the sixth factor loaded poorly and had an eigenvalue less than 1. Therefore, we determined the five-factor solution best represented the data.

Following the iterative deletion of five items not meeting the .50/.30 criteria, the final, five-factor solution collectively explained 60.53% of the variance. The first factor was labeled *Norms* ($M = 2.22$, $SD = 1.15$, $\omega = .855$). This dimension consisted of five items relating to an instructor's assumed expectations for classroom decorum. The second factor was labeled *Deadlines* ($M = 3.36$, $SD = 1.46$, $\omega = .854$). This dimension included three items related to the instructor's enforcement of policies for accepting late work. The third factor was labeled *Testing* ($M = 3.91$, $SD = 1.62$, $\omega = .782$). This dimension included three items related to the extent the instructor enforced policies about cheating on tests or exams. The fourth factor was labeled *Technology* ($M = 2.49$, $SD = 1.44$, $\omega = .880$). This dimension included three items related to the way the instructor enforced

**Table 5.** Rotated factor structure of regulatory strictness.

| Survey item | Factor | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| *Norms* | | | | | |
| 1. Forbade students from using the restroom during class. | **.94** | –.03 | .00 | –.04 | –.01 |
| 2. Enforced classroom rules that were not directly stated in the syllabus. | **.71** | .03 | .04 | –.04 | .02 |
| 3. Restricted how students took notes during class. | **.71** | .02 | –.02 | .20 | –.08 |
| 4. Did not allow students to talk informally with one another during class. | **.56** | .07 | –.04 | .09 | .09 |
| 5. Banned students from eating during class. | **.50** | .12 | .05 | .13 | .11 |
| *Deadlines* | | | | | |
| 6. Refused to accept late work. | –.08 | **.97** | –.01 | .04 | –.04 |
| 7. Would not grade an assignment if it was turned in late. | .13 | **.75** | .01 | –.03 | –.01 |
| 8. Declined to extend assignment deadlines. | .11 | **.63** | .07 | –.06 | .13 |
| *Testing* | | | | | |
| 9. Took steps to ensure students could not cheat on tests or exams. | –.04 | –.03 | **.92** | .01 | –.01 |
| 10. Made students take tests at the same time to prevent cheating. | .00 | .06 | **.68** | .01 | .01 |
| 11. Used technology (e.g., software, camera) to closely monitor individual behavior when taking tests. | .22 | .04 | **.54** | –.02 | .02 |
| *Technology* | | | | | |
| 12. Punished students for inappropriately using of technology in class. | –.01 | –.03 | .05 | **.88** | –.03 |
| 13. Controlled how students used technology (e.g., phones, laptops; cameras) in class. | .00 | –.03 | –.01 | **.83** | .08 |
| 14. Publicly called out students for using personal devices (e.g., phones, laptops) in class. | .03 | .08 | –.02 | **.81** | –.03 |
| 15. Did not allow students to use personal devices (e.g., phones, laptops) in class. | .19 | .03 | –.02 | **.55** | .06 |
| *Attendance* | | | | | |
| 16. Required students to attend class meetings. | .01 | –.03 | –.08 | –.02 | **.70** |
| 17. Needed documentation to excuse student absences. | –.05 | .03 | .13 | .08 | **.61** |
| SS loading | 2.84 | 2.06 | 1.69 | 2.71 | 0.96 |
| Variance accounted for | 16.75% | 12.20% | 10.00% | 15.94% | 5.64% |

Note. SS loading = rotated sum of squares loading. Bold factor coefficients show acceptable factor loadings for the corresponding items and factors. Items that are not in bold did not load on the corresponding factor.

policies controlling classroom technology. The fifth factor was labeled *Attendance* ($M$ = 4.75, $SD$ = 1.63, α = .605). This dimension included two items related to the way the instructor enforced classroom attendance policies. The comprehensive scale demonstrated high internal reliability of ω = .849. Using cutoff criteria suggested by Hu and Bentler (1999), the final five-factor solution fit the data well: $\chi^2(61) = 122.677$, $p < .001$; RMSEA = .051 (.038, .064); CFI = .980, TLI = .955, SRMR = .019. Factors loadings for individual items in the final model are presented in Table 5.

### *Interactive strictness*

The scree plot suggested that two factors best fit the 15 items for interactive strictness. Contrarily, the PA suggested a three-factor model. Thus, we examined models with solutions ranging from one to three factors. Both the KMO measure (0.93) and Bartlett's test [$\chi^2 = 5199.433$ (105), $p < .001$] were acceptable. The results demonstrated worsening of model fit when moving from the three-factor model to the two and one factor models; we determined that the three-factor solution best represented the data.

After the iterative deletion of three items not meeting the .50/.30 criteria, the final three-factor solution collectively explained 74.33% of the variance. The first factor consisted of two items and was labeled *Excuses* ($M$ = 2.67, $SD$ = 1.46, α = .913). The two items

**Table 6.** Rotated factor structure of interactive strictness.

| Survey item | Factor | | |
|---|---|---|---|
| | 1 | 2 | 3 |
| *Excuses* | | | |
| 1. Did not listen to student excuses. | **1.01** | .02 | −.04 |
| 2. Was not willing to listen to student excuses. | **.74** | .01 | .17 |
| *Rudeness* | | | |
| 3. Was mean to students who did not follow rules or policies. | −.01 | **.95** | .01 |
| 4. Took frustration out on students when they did not follow a rule or policy. | −.07 | **.92** | .07 |
| 5. Spoke to students with a stern tone. | .07 | **.80** | −.14 |
| 6. Made disparaging comments toward students who did not follow classroom rules. | .06 | **.77** | .06 |
| 7. Belittled students if they did not follow classroom rules. | .18 | **.59** | .14 |
| *Availability* | | | |
| 8. Was unavailable to help students solve problems. | .02 | −.02 | **.94** |
| 9. Was unavailable to help students complete tasks. | .00 | .02 | **.90** |
| 10. Was not available for additional help outside of set class time. | −.04 | −.01 | **.79** |
| 11. Showed no interest in helping students. | .06 | .21 | **.70** |
| 12. Went out of their way to make themselves available to students. [R] | .17 | −.07 | **.50** |
| SS loading | 1.87 | 3.64 | 3.41 |
| Variance accounted for | 15.57% | 30.34% | 28.42% |

Note. SS loading = rotated sum of squares loading. [R] = item is reverse-coded. Bold factor coefficients show acceptable factor loadings for the corresponding items and factors. Items that are not in bold did not load on the corresponding factor.

were related to the instructor's willingness to listen to students' excuses for failing to follow or breaking a policy. The second factor was labeled *Rudeness* ($M = 2.25$, $SD = 1.27$, $\omega = .931$). This dimension included five items related to the way an instructor treated students after they failed to follow or broke a policy. The third factor consisted of five items and was labeled *Availability* ($M = 2.19$, $SD = 1.17$, $\omega = .900$). The items were related to the instructor's inflexibility with their time to help students with tasks or problems. The comprehensive scale demonstrated high internal reliability of $\omega = .945$. The final three-factor solution fit the data adequately: $\chi^2(33) = 167.225$, $p < .001$; RMSEA = .102 (.087, .118); CFI = .970, TLI = .940, SRMR = .021. Factors loadings for individual items in the final model are presented in Table 6.

## Correlation analyses

The remaining hypotheses aimed to provide convergent validity evidence for each proposed strictness instrument. Correlations between the combined version of each instrument, individual factors, and the remaining study variables are provided in Table 7.

H1 predicted that each proposed measure would be positively associated with the strictness inventory from Wubbels et al.'s (1985) QTI. Table 7 confirmed significant, positive correlations between the holistic scales and individual factors for all three instruments. H1 was supported.

H2 predicted that each proposed measure would be positively associated with the admonishing behavior inventory from Wubbels et al.'s (1985) QTI. Table 7 confirmed significant, positive correlations between the holistic scales and individual factors for all three instruments, with one exception. The correlation between the *Attendance* dimension of regulatory strictness and admonishing behavior was not significant ($r = 0.07$, $p = .073$). H2 was partially supported.

**Table 7.** Correlations (one-tailed).

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. QTI—Strictness | – | | | | | | | | | | | | | | | |
| 2. QTI—Admonishing | 0.55 | – | | | | | | | | | | | | | | |
| 3. Standards | 0.38 | 0.32 | – | | | | | | | | | | | | | |
| 4. Rounding | 0.35 | 0.25 | 0.32 | – | | | | | | | | | | | | |
| 5. Harshness | 0.56 | 0.43 | 0.51 | 0.43 | – | | | | | | | | | | | |
| 6. Ambiguity | 0.43 | 0.41 | 0.66 | 0.32 | 0.59 | – | | | | | | | | | | |
| 7. Normative | 0.37 | 0.49 | 0.39 | 0.35 | 0.48 | 0.44 | – | | | | | | | | | |
| 8. Deadline | 0.47 | 0.38 | 0.37 | 0.38 | 0.53 | 0.49 | 0.43 | – | | | | | | | | |
| 9. Testing | 0.38 | 0.20 | 0.28 | 0.16 | 0.30 | 0.31 | 0.24 | 0.32 | – | | | | | | | |
| 10. Technological | 0.38 | 0.44 | 0.28 | 0.31 | 0.43 | 0.37 | 0.63 | 0.37 | 0.20 | – | | | | | | |
| 11. Attendance | 0.20 | 0.07* | 0.10* | 0.15 | 0.24 | 0.09* | 0.13 | 0.10* | 0.15 | 0.22 | – | | | | | |
| 12. Excuses | 0.53 | 0.46 | 0.36 | 0.27 | 0.47 | 0.46 | 0.49 | 0.50 | 0.26 | 0.38 | 0.11* | – | | | | |
| 13. Rudeness | 0.56 | 0.68 | 0.38 | 0.32 | 0.51 | 0.48 | 0.64 | 0.44 | 0.27 | 0.58 | 0.10* | 0.64 | – | | | |
| 14. Availability | 0.52 | 0.60 | 0.43 | 0.26 | 0.52 | 0.55 | 0.60 | 0.49 | 0.23 | 0.44 | 0.03* | 0.62 | 0.76 | – | | |
| 15. Total Eval. | 0.55 | 0.44 | 0.81 | 0.65 | 0.81 | 0.82 | 0.54 | 0.56 | 0.33 | 0.45 | 0.19 | 0.51 | 0.55 | 0.56 | – | |
| 16. Total Reg. | 0.55 | 0.51 | 0.44 | 0.42 | 0.61 | 0.54 | 0.80 | 0.67 | 0.57 | 0.79 | 0.40 | 0.55 | 0.65 | 0.58 | 0.65 | – |
| 17. Total Inter. | 0.60 | 0.67 | 0.44 | 0.32 | 0.56 | 0.56 | 0.66 | 0.53 | 0.28 | 0.54 | 0.08* | 0.78 | 0.93 | 0.92 | 0.61 | 0.67 |

Note. All correlations are significant at $p < .01$ unless marked with an *.

## Study three: CFA and concurrent validity

Study 3 provides further validity evidence by subjecting the measures forwarded in Study 2 to CFA procedures. In addition, this study also establishes concurrent validity through expected relationships between each form of instructor strictness, instructor caring, and cognitive flexibility.

### *Instructor caring*

Concurrent validity is established when an explanatory variable is related to a different response variable in an expected direction. Despite the treatment of strictness as an inherently value-free construct, it is likely that students will hold a negative connotation of the concept in the absence of other factors (e.g., the instructor's framing of a policy; students' attributions for the inclusion of a policy; Poplin et al., 2011). This is in stark contrast to research concerning instructor caring, or perceived instructor concern for the well-being of students (McCroskey, 1998; Teven & McCroskey, 1997). Research suggests that instructors who harshly evaluate their students' work, routinely enforce classroom policies and procedures, or come across as mean, rude, or aggressive are likely to be seen as less caring (Banfield et al., 2006; Thweatt & McCroskey, 1998). Thus, the three measures of instructor strictness should be negatively related to students' perceptions of instructor caring.

### *Cognitive flexibility*

Flexible communicators change and adapt their behavior to be effective within their social situation (Martin & Rubin, 1994, p. 173). Flexible individuals recognize that they have a variety of communicative alternatives that they can choose from, and they demonstrate competence by adapting and applying the behaviors that are most appropriate for the context of their interactions (Spitzberg & Cupach, 1984, 1989). However, as Martin and Rubin (1995) note, before one can be flexible in a given social situation, they must be cognitively flexible. Cognitive flexibility is "a person's (a) *awareness* that in any given situation there are options and alternatives available, (b) *willingness* to be flexible and adapt to the situation, and (c) *self-efficacy* or belief that one has the ability to be flexible" (Martin & Anderson, 1998, p. 1). Given that strictness is conceptually rooted in the concept of inflexibility, we expect that students who perceive instructors as stricter also perceive instructors as less aware, willing, and capable of changing their behavior. Accordingly, the three measures of instructor strictness should be negatively related to students' perceptions of cognitive flexibility. The following hypotheses were posed to guide study 3:

H3: The factor structures for (a) evaluative, (b) regulatory, and (c) interactive strictness proposed in study 2 will fit the data well.

H4: Each proposed measure will be negatively associated with instructor caring.

H5: Each proposed measure will be negatively associated with cognitive flexibility.

### Participants and procedures

Participants ($N = 528$) included undergraduate students enrolled in communication courses at a large, Southeastern university (151 men, 28.60%; 376 women, 71.21%; 1 did not report; 0.19%). The participants' ages ranged from 18 to 49 ($M = 20.10$, $SD = 3.03$). Students' class ranks varied: first-year students (36.93%), sophomore (17.42%), junior (29.36%), seniors (14.96%), fifth-year senior or beyond (0.19%), graduate student (0.19%), and unreported (0.95%). They reported their ethnicity as White/Caucasian ($n = 432$; 81.82%), Black or African American ($n = 31$; 5.87%), Asian or Pacific Islander ($n = 27$; 5.11%), Hispanic or Latino/Latina ($n = 20$; 3.79%), Native American or American Indian ($n = 1$; 0.19%), Bi-racial or Mixed ($n = 10$; 1.89%), Other ($n = 6$; 1.14%), and unreported ($n = 1$; 0.19%). The procedures for this study were identical to those described in study 2.

### Measures

#### Evaluative strictness

Students responded to the 12 items assessing evaluative strictness developed in study 2. Items referenced the stem "My instructor … " and were drafted using a Likert-type scale ranging from *Strongly Disagree* (1) to *Strongly Agree* (7). Higher scores indicated greater perceived levels of inflexible adherence to the respective dimension (i.e., greater strictness): Standards ($M = 3.41$, $SD = 1.42$, $\omega = .796$), Rounding ($M = 4.11$, $SD = 1.41$, $\omega = .877$), Harshness ($M = 3.26$, $SD = 1.38$, $\omega = .820$), Ambiguity ($M = 2.79$, $SD = 1.46$, $\omega = .830$), Evaluative Strictness ($M = 3.39$, $SD = 1.11$, $\omega = .884$).

#### Regulatory strictness

Students responded to the 17 items assessing regulatory strictness developed in study 2. Items referenced the stem "My instructor … " and were drafted using a Likert-type scale ranging from *Strongly Disagree* (1) to *Strongly Agree* (7). Higher scores indicated greater perceived levels of regulatory strictness: Norms ($M = 2.24$, $SD = 1.06$, $\omega = .814$), Deadlines ($M = 3.61$, $SD = 1.50$, $\omega = .861$), Testing ($M = 4.15$, $SD = 1.59$, $\omega = .764$), Technology ($M = 2.58$, $SD = 1.49$, $\omega = .882$), Attendance ($M = 4.89$, $SD = 1.57$, $\alpha = .564$), Regulatory Strictness ($M = 3.20$, $SD = 0.88$, $\omega = .823$).

#### Interactive strictness

Students responded to the 12 items assessing regulatory strictness developed in study 2. Items referenced the stem "My instructor … " and were drafted using a Likert-type scale ranging from *Strongly Disagree* (1) to *Strongly Agree* (7). Higher scores indicated greater perceived levels of interactive strictness: Excuses ($M = 2.73$, $SD = 1.50$, $\alpha = .902$), Rudeness ($M = 2.20$, $SD = 1.21$, $\omega = .933$), Availability ($M = 2.14$, $SD = 1.13$, $\omega = .895$), Interactive Strictness ($M = 2.27$, $SD = 1.10$, $\omega = .941$).

#### Caring

Instructor caring was operationalized using Teven and McCroskey's (1997) Caring Scale. This instrument consists of six items asking students to report perceptions of instructor caring using a semantic differential format with adjectives placed at opposite ends of a 7-

point scale (e.g., "Not understanding—Understanding"). Higher scores indicate greater perceived levels of caring ($M = 5.73$, $SD = 1.10$, $\omega = .873$).

### Cognitive flexibility

Cognitive flexibility was measured using a modified version of Martin and Rubin's (1995) *Cognitive Flexibility Scale*. This scale included 12 items (e.g., "My instructor is able to act appropriately in any given situation", "My instructor is willing to listen and consider alternatives for handling a problem") collected using a Likert-type scale ranging from 1 (*Almost Never True*) to 7 (*Almost Always True*). Higher scores indicated greater perceived levels of cognitive flexibility ($M = 5.33$, $SD = 0.98$, $\omega = .782$).

## Results

### CFAs

Correlations between the combined version of each instrument, individual factors, and the study 3 variables are provided in Table 8.

To test the hypotheses and provide structural and construct validity for the scales, individual CFAs with robust maximum likelihood estimation were conducted via the *lavaan* package (Rosseel, 2012) in the free statistical software RStudio (Version 1.4.1717; RStudio Team, 2021); all codes are provided in the online appendix.[2] Consistent with the procedures outlined in study 2, model fit was evaluated based on criteria from Kline (2016) and Hu and Bentler (1999): (1) chi-square goodness-of-fit, (2) root mean square error of approximation (RMSEA) <.08, (3) standardized root mean square residual (SRMR) <.08, (4) Tucker–Lewis index (TLI) >.90, (5) comparative fit index (CFI) >.90.

Prior to running the analyses, a decision was made to drop the *Attendance* dimension from the measure of regulatory strictness. This dimension had poor reliability and was the only subscale that was not significantly correlated with either the other strictness instruments or validity measures in both study 2 and study 3 (see Tables 7 and 8). Therefore, analyses included a four-factor model of regulatory strictness.

For evaluative strictness, global fit indices indicated that the four-factor structure fit the data reasonably well: $\chi^2$ (48) = 114.16, $p < .000$, scaling correction factor = 1.215, RMSEA = .056 [90% CI: .043, .070], SRMR = .040, TLI = .968, CFI = .977. All 12 items loaded significantly on their respective factors (loadings ranged from 0.43 to 0.93).[3] An examination of the normalized residuals also indicated that the model fit well at a local level.

For regulatory strictness, global fit indices indicated that the four-factor structure fit the data reasonably well: $\chi^2$ (84) = 181.037, $p < .000$, scaling correction factor = 1.322, RMSEA = .054 [90% CI: .043, .065], SRMR = .054, TLI = .955, CFI = .964. All 15 items loaded significantly on their respective factors (ranging from 0.57 to 0.93). However, an examination of the normalized residuals revealed several instances of local misfit for one item (i.e., "Uses technology (e.g., software, cameras) to closely monitor individual behavior when taking tests"). Local misfit can be considered significant when residuals are greater than +/- 2.58 (2 standard deviations; Bowman & Goodboy, 2020; Goodboy & Kline, 2017). This poor local fit was likely caused by the shared reference to technology between this item and the *Technology* subscale. Thus, the model was respecified with this item removed. In this iteration, global fit indices indicated that the revised four-factor

**Table 8.** Correlations (one-tailed).

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Caring | – | | | | | | | | | | | | | | | |
| 2. Cog. Flexibility | 0.50 | – | | | | | | | | | | | | | | |
| 3. Standards | −0.32 | −0.35 | – | | | | | | | | | | | | | |
| 4. Rounding | −0.35 | −0.30 | 0.34 | – | | | | | | | | | | | | |
| 5. Harshness | −0.42 | −0.43 | 0.49 | 0.44 | – | | | | | | | | | | | |
| 6. Ambiguity | −0.47 | −0.52 | 0.62 | 0.38 | 0.62 | – | | | | | | | | | | |
| 7. Normative | −0.36 | −0.53 | 0.37 | 0.25 | 0.39 | 0.48 | – | | | | | | | | | |
| 8. Deadline | −0.38 | −0.33 | 0.39 | 0.43 | 0.46 | 0.39 | 0.29 | – | | | | | | | | |
| 9. Testing | −0.16 | −0.19 | 0.28 | 0.19 | 0.24 | 0.21 | 0.16 | 0.27 | – | | | | | | | |
| 10. Technological | −0.26 | −0.28 | 0.30 | 0.24 | 0.35 | 0.40 | 0.52 | 0.27 | 0.19 | – | | | | | | |
| 11. Attendance | −0.09* | 0.06* | 0.13 | 0.20 | 0.19 | 0.14 | 0.08* | 0.22 | 0.12 | 0.17 | – | | | | | |
| 12. Excuses | −0.43 | −0.55 | 0.43 | 0.37 | 0.48 | 0.54 | 0.52 | 0.47 | 0.19 | 0.41 | 0.12 | – | | | | |
| 13. Rudeness | −0.45 | −0.60 | 0.39 | 0.30 | 0.46 | 0.52 | 0.65 | 0.40 | 0.22 | 0.50 | 0.06* | 0.64 | – | | | |
| 14. Availability | −0.50 | −0.69 | 0.32 | 0.31 | 0.44 | 0.51 | 0.56 | 0.32 | 0.17 | 0.36 | 0.03* | 0.56 | 0.74 | – | | |
| 15. Total Eval. | −0.50 | −0.51 | 0.79 | 0.69 | 0.81 | 0.84 | 0.48 | 0.53 | 0.29 | 0.41 | 0.21 | 0.58 | 0.54 | 0.51 | – | |
| 16. Total Reg. | −0.40 | −0.43 | 0.46 | 0.40 | 0.52 | 0.52 | 0.72 | 0.63 | 0.55 | 0.74 | 0.40 | 0.55 | 0.63 | 0.50 | 0.61 | – |
| 17. Total Inter. | −0.54 | −0.70 | 0.42 | 0.36 | 0.51 | 0.58 | 0.66 | 0.44 | 0.21 | 0.48 | 0.07* | 0.77 | 0.93 | 0.89 | 0.60 | 0.62 |

Note. All correlations are significant at $p < .01$ unless marked with an *.

structure fit the data well: $\chi^2$ (71) = 137.349, $p < .000$, scaling correction factor = 1.350, RMSEA = .049 [90% CI: .036, .061], SRMR = .042, TLI = .966, CFI = .973. The 14 items again loaded significantly on their respective factors ($\omega = .782$), and the normalized residuals indicated that the model fit well at a local level.

For interactive strictness, global fit indices indicated that the three-factor structure fit the data well: $\chi^2$ (51) = 121.608, $p < .000$, scaling correction factor = 1.533, RMSEA = .063 [90% CI: .049, .078], SRMR = .028, TLI = .976, CFI = .982. All 12 items loaded significantly on their respective factors (ranging from 0.45 to 0.96). An examination of the normalized residuals also indicated that the model fit relatively well at a local level. Collectively, H3 was partially supported.

### *Correlational analyses*

H4 predicted that each proposed measure would be negatively associated with instructor caring. Table 8 confirmed significant, negative correlations between the holistic scales and individual factors for all three instruments. Thus, H4 was supported.

H5 predicted that each proposed measure would be negatively associated with cognitive flexibility. Table 8 confirmed significant, negative correlations between the holistic scales and individual factors for all three instruments. Thus, H5 was supported.

## Discussion

The idea that an instructor can be *strict* is pervasive throughout much of popular culture and generally accepted as a common practice. Parents, teachers, students, and many other individuals who inhabit instructional contexts are likely to be familiar with the term and its implications. However, the little research that has been conducted concerning strict instruction fails to provide a comprehensive definition of the concept or capture the wide domain of strictness through existing measures (e.g., Jiang et al., 2021). This three-part study presented a more inclusive and encompassing conceptualization and operationalization of instructor strictness that fills existing knowledge gaps and expands thinking about how the enforcement of classroom policies and procedures contributes to students' instructional experiences.

Drawing on the conceptual definition provided by Tatum and Frey (2021), study 1 used open-ended student responses to better understand what instructor strictness looks like in the college classroom. A thematic analysis of the responses indicated that students perceived instructors to remain steadfast in their policies and procedures in three general areas: evaluation, classroom regulation (through explicit policies and norms), and interaction. These findings were then integrated into a pool of items representing the different types of strictness, which was further confirmed by a focus group of undergraduate students.

The responses build and enhance upon the work done by Wubbels et al. (1985) in understanding strict instruction as presented through the QTI. Students' experiences with strict instruction were broadened to include inflexible enforcement of attendance, technology, testing, and late work policies, classroom norms not explicitly stated in the syllabus (e.g., informal interactions, grade rounding behaviors, ambiguity), and particularly high standards related to the amount of work students must complete within the course. Particularly important is that students identified the way that an instructor interpersonally

treated them (e.g., not listening to excuses, rudeness, a lack of availability) as a component of strict behavior. Within the QTI, this idea appeared to be captured not by strictness but by another construct in the form of admonishing behavior. Strictness is concerned with the rigid adherence to policies and procedures, and students clearly articulated the relational component of this enforcement in their narratives. Taken together, the conclusions from the first study expand the current understanding of strictness and present opportunities to analyze how this instructional behavior influences classroom experiences in areas beyond those identified within the QTI and similar instruments.

The results from study 2 provided factorial and convergent validity for the evaluative, regulatory, and interactive strictness measures. Each scale consists of several factors that represent the variety of policies or procedures that an instructor realistically implements in a classroom setting. In addition, both the individual subscales and the holistic scales were reliable. Convergent validity was demonstrated through moderately strong, positive relationships between all strictness measures and both the strict and admonishing behavior dimensions from the QTI. Thus, the instructor strictness scales were related to but not redundant with existing scales of the same construct; the measures are similar but not isomorphic.

Study 3 provided construct and concurrent validity by subjecting the new measures to a series of CFAs and investigating relationships between instructor caring and cognitive flexibility. The four-factor and three-factor models for evaluative and interactive strictness that were proposed in study 2, respectively, fit the data well. However, the five-factor model for regulatory strictness required revision.

Specifically, the *Attendance* dimension was dropped from the measure. In addition to poor reliability and nonsignificant correlations with theorized constructs, an argument could be made that the items for this dimension are not truly representative of inflexibility or a strict adherence to a policy ("Required students to attend class meetings," "Needed documentation to excuse student absences"). Furthermore, many students and instructors may not see attendance as a primary responsibility necessary for facilitating effective instruction. Instead, perhaps many instructors exclude attendance policies or rely on scripted language mandated by university administration for tuition and enrollment purposes (e.g., Title IV). Said differently, students might perceive that requirements to attend class or provide documentation are enforced at an administrative level rather than through an individual instructor's actions. This is also reflected in the mean for the attendance dimension in both study 2 and study 3, which was well above the mean value and appears to suggest that attendance was enforced more universally (i.e., as an expectation of the institution) than many of the other policies. Second, one item was dropped from the *Testing* dimension due to poor local fit. This resulted in a better fitting and more conceptually sound model, as the item in question shared commonalities in wording with the *Technology* dimension.

Study 3 also provided support for construct validity through hypothesized relationships between each form of strictness and instructor caring and cognitive flexibility. Concurrent validity was demonstrated through the significant relationships between each measure of strictness and perceived instructor caring. Instructor caring is based on empathy (i.e., the ability to identify with another's feelings or situation; Stiff et al., 1988), understanding (i.e., the ability to comprehend another's needs, ideas, and feelings; Cahn & Shulman, 1984), and responsiveness (i.e., having a sensitivity toward others by being other-orientated;

Thomas et al., 1994). Instructors who are caring can adjust their behavior to show support to students when they need it. Likewise, cognitively flexible communicators have the knowledge, motivation, and capability to adjust their behavior given the constraints of a situation. Thus, cognitive flexibility is generally treated as a central component of competent communication (Duran, 1983; Martin & Rubin, 1994), and inflexibility is likely to be seen as incompetent or ineffective. In isolation, we should expect these variables to be related, and these results were confirmed by the study.

However, we believe these results potentially highlight how the uniqueness of the instructional setting may impact the relationship between strictness and other outcomes. For example, there are clear differences in broad, cultural expectations for instructors to enact strict and authoritative behaviors (Biggs & Watkins, 2001; Jiang et al., 2021; Zhang & Zhu, 2008). These differences may also manifest at geographical, university, departmental, or disciplinary levels; institutional expectations for student behavior or involvement, or even standards of student scholarship, may impact how strict instructor behavior is differentially interpreted. Thus, we recognize the call by Mazer and Graham (2015) to assess measures across culturally diverse samples. Future research might seek to confirm measurement invariance of each strictness instrument across groups of students who potentially respond to the items in different ways, providing a stringent test of the factor structures proposed in this study.

Several student narratives in study 1 also hinted at the possibility that strictness may be tolerated, or perhaps even encouraged, in an instructional setting (e.g., to keep students safe in a lab; to treat all students fairly). We speculate that effects of strictness may be mitigated depending on students' perceptions of an instructor's intent when enforcing a policy or procedure, similar to how scholars have conceptualized aggressive instructor behavior (e.g., Martin et al., 2010). Aggressive communication can take both constructive and destructive forms, depending on how it is received by the student. As Avtgis (2016) stated, "it appears as if students' perceptions determine whether instructor communication behavior is interpreted as argumentative (constructive) or verbally aggressive (destructive) in nature" (p. 268).

Said differently, if a student perceives the purpose of an instructor's perceived inflexibility is to ultimately benefit them (i.e., there is a good reason this policy or procedure is in place) or that it is communicated without malice or aggression, they may evaluate strictness positively. Indeed, an instructor has the capacity to be both strict and supportive simultaneously. For example, an instructor can identify with a student's feelings of disappointment about a grade (i.e., empathy), even though he or she was the one who distributed the grade using stringent criteria (i.e., strictness), or an instructor could recognize why an assignment was submitted after a deadline had passed (i.e., understanding) and simultaneously express the desire to hold students to the same standards as potential future employers (i.e., strictness). This potentially opens the door for new and important work concerning how perceptions of strictness interact with other instructional variables to influence classroom outcomes.

Accordingly, one avenue ripe for exploration is the relationship between strictness and caring demonstrated in this study. Research on parenting behaviors has demonstrated that children of authoritative parents—who are responsive (i.e., accepting, supportive, caring) and demanding (i.e., controlling, strict)—receive better academic grades (Radziszewska et al., 1996), tend to be more academically competent (Steinberg et al., 2006), and

are less likely to engage in antisocial behavior (Steinberg, 2001). Essentially, the authoritative parent "affirms the child's present qualities, but also sets standards for future conduct" (Baumrind, 1971, p. 261). Considering the increasingly close relationships college students have with their parents (Elam et al., 2007; Frey & Tatum, 2016), perhaps the same stylistic patterns of caring and strictness may be effective in promoting positive student outcomes in the classroom.

A final important conclusion from this three-part study concerns the assortment of individual factors that emerged representing the three patterns of strictness. Since strictness occurs in relation to policies or procedures, having several factors that represent the variety of places where rules can be applied aligns with the conceptual definition. The difficulty lies in understanding the best strategy for implementing the measures to assess classroom practices; researchers might choose to isolate specific policies depending on their research questions and hypotheses.

To illustrate, several studies have investigated students' reactions to classroom messages in the form of technology policies. This research shows clear differences in student responses depending on the framing of the policy; students do not like their autonomy restricted by classroom rules in general, but they appear to respond more favorably when those rules and communicated in ways that are encouraging (Finn & Ledbetter, 2013, 2014), fair, and less controlling (Frey et al., 2021). However, the classroom context presents a distinct scenario in which students know their options to restore autonomy are limited out of a need for academic survival (i.e., their status or grades may be affected depending on their behaviors; Horan et al., 2012). Perhaps Finn and Ledbetter (2013) summarized it best when they stated that it "is likely a combination of (a) the policy and (b) the way the instructor communicates and enforces classroom policies and procedures that influences students' perceptions" (p. 39). Instructor strictness presents an opportunity for researchers to better investigate both the message and the enforcement of it to paint a more nuanced picture of how instructors manage their classrooms through rules, policies, and norms.

## Limitations and future research

One important limitation of the current study is the lack of additional validation to support the measures. Divergent validity is particularly important in this case. There are clear conceptual similarities between instructor strictness, misbehaviors, and classroom justice, and steps should be taken to ensure instruments are truly assessing different constructs. Another form of structural validation that may be necessary for strictness involves comparing the first-order factor structures reported here to respecified models with second-order factor structures, depicting the respective dimensions for each scale as the lower order factors and the pattern of strictness (i.e., evaluative, regulatory, interactive) as the higher order factors (Chen et al., 2005). Byrne (2005) posed four questions related to the adoption of first-order versus second-order models: (1) Does the higher-order structure represent a well-fitting model? (2) Is the discrepancy between model fit minimal? (3) Are the correlations between the latent factors substantial? and (4) Is there theoretical justification to consider a higher order construct? In the current study, one could argue that the correlations between the latent factors for each measure were not large enough to justify the inclusion of a second-order model; in

Byrne's (2005) study, correlations between latent factors ranged from 0.77 to 0.93, much stronger than most of those observed for the current research. However, there is theoretical justification (i.e., the results of study 1) to suggest that each latent factor in the first-order structure should be represented by second-order factors (i.e., each pattern of strictness). Researchers interested in strictness are encouraged to evaluate competing models to determine whether the measures are best represented as multidimensional or unidimensional constructs.

Second, this study is limited in that the relationships between strictness and learning were not assessed. Learning outcomes are the ultimate goal for instructional communication researchers (Clark, 2002), and future research should seek to examine the impact that students' perceptions of strict instructor behaviors have on their reported learning outcomes. Researchers should anticipate a negative relationship between strictness and students' learning outcomes, yet they should also remain mindful of the possibility that the impact of strictness on students' motivation, affective learning, cognitive learning, and satisfaction (among other outcome variables) is conditional based on other factors.

Third, this study is limited in its reliance solely on self-report data. There is clear value in using self-report measures to better understand students' experiences, but this approach is limited until it can be complemented with other validation approaches. Future research should seek to provide additional validation by building a multitrait-multimethod matrix (Campbell & Fiske, 1959) that assesses strictness in a variety of ways. For example, students' perceptions of instructor strictness may differ widely from instructors' perceptions of their own strict behavior. Or, perhaps observational research would reveal that instructors rarely engage in the low-inference behaviors that make up each respective scale. Finally, a short-form, high inference measure of strictness that reflects the inherent characterizations of strict instructors would be applicable in a wider variety of settings, contexts, and situations. Evidence that any of the above approaches are strongly, and positively related to the advocated measures of strictness would provide important evidence of convergent validity and ensure the scales are performing as intended.

Ultimately, strictness may not be the damaging, disreputable teaching behavior that is often portrayed in popular media and press. Looking at the items that make up each scale, it would be irresponsible not to suggest that it can, of course, have consequences if wielded inappropriately. Questions remain regarding how an instructor's rigid enforcement (or renunciation) of rules, policies, and norms influences learning and achievement, what short- and long-term effects inflexible instruction might have throughout an academic term, how strictness impacts instructor–student relationships, and whether strict behavior can potentially motivate students to higher levels of achievement. Research that seeks to answer such questions can provide a fresh theoretical perspective to investigate the light side of seemingly dark instructor behaviors, and we welcome new research that accepts this challenge.

## Notes

1. https://osf.io/3hyds/?view_only=5878be0e5feb45d690169a17380d9747.
2. https://osf.io/3hyds/?view_only=5878be0e5feb45d690169a17380d9747.
3. Unstandardized and standardized CFA loadings are available via the online appendix: https://osf.io/3hyds/?view_only=5878be0e5feb45d690169a17380d9747.

## Disclosure statement

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

## ORCID

*T. Kody Frey* http://orcid.org/0000-0003-1001-8821
*Nicholas T. Tatum* http://orcid.org/0000-0002-1688-1717

## References

Avtgis, T. A. (2016). Instructor use of aggressive communication. In P. L. Witt (Ed.), *Communication and learning* (pp. 259–286). De Gruyter Mouton.

Banfield, S. R., Richmond, V. P., & McCroskey, J. C. (2006). The effect of teacher misbehaviors on teacher credibility and affect for the teacher. *Communication Education*, *55*(1), 63–72. https://doi.org/10.1080/03634520500343400

Baumrind, D. (1971). Current patterns of parental authority. *Developmental Psychology*, *4*(1), 1–103. https://doi.org/10.1037/h0030372

Biggs, J. B., & Watkins, D. A. (2001). Insights into teaching the Chinese learner. In D. A. Watkins & J. B. Biggs (Eds.), *Teaching the Chinese learner: Psychological and pedagogical perspectives* (pp. 277–298). Comparative Education Research Center.

Bowman, N. D., & Goodboy, A. K. (2020). Evolving considerations and empirical approaches to construct validity in communication science. *Annals of the International Communication Association*, *44*(3), 219–234. https://doi.org/10.1080/23808985.2020.1792791

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Byrne, B. M. (2005). Factor analytic models: Viewing the structure of an assessment instrument from three perspectives. *Journal of Personality Assessment*, *85*(1), 17–32. https://doi.org/10.1207/s15327752jpa8501_02

Cahn, D. D., & Shulman, G. M. (1984). The perceived understanding instrument. *Communication Research Reports*, *1*(1), 122–125.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81–105. https://doi.org/10.1037/h0046016

Carpenter, S. (2018). Ten steps in scale development and reporting: A guide for researchers. *Communication Methods and Measures*, *12*(1), 25–44. https://doi.org/10.1080/19312458.2017.1396583

Carr, S. (2014, December). How strict is too strict? The backlash against no-excuses discipline in high school. *The Atlantic*. https://www.theatlantic.com/magazine/archive/2014/12/how-strict-is-too-strict/382228/

Chen, F. F., Sousa, K. H., & West, S. G. (2005). Teacher's corner: Testing measurement invariance of second-order factor models. *Structural Equation Modeling: A Multidisciplinary Journal*, *12*(3), 471–492. https://doi.org/10.1207/s15328007sem1203_7

Chory-Assad, R. M., & Paulsel, M. L. (2004). Classroom justice: Student aggression and resistance as reactions to perceived unfairness. *Communication Education*, *53*(3), 253–273. https://doi.org/10.1080/0363452042000265189

Clark, R. A. (2002). Learning outcomes: The bottom line. *Communication Education*, *51*(4), 396–404. https://doi.org/10.1080/03634520216531

Coll, R. K., Taylor, N., & Fisher, D. L. (2002). An application of the questionnaire on teacher interaction and college and university classroom environment inventory in a multicultural tertiary context. *Research in Science & Technological Education*, *20*(2), 165–183. https://doi.org/10.1080/0263514022000030435

Darling-Hammond, L., & Hyler, M. E. (2020). Preparing educators for the time of COVID … and beyond. *European Journal of Teacher Education*, *43*(4), 457–465. https://doi.org/10.1080/02619768.2020.1816961

Den Brok, P., & Brekelmans, M. (2001). Internal quality assurance in education: An instrument for the assessment of interpersonal teaching skills. *Kvalita Inovacia Prosperita*, *5*(2), 22–29. https://qip-journal.eu/files/2001/2/KIP-2-2001.pdf#page=32

DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed.). Sage.

Duran, R. L. (1983). Communicative adaptability: A measure of social communicative competence. *Communication Quarterly*, *31*(4), 320–326. https://doi.org/10.1080/01463378309369521

Elam, C., Stratton, T., & Gibson, D. D. (2007). Welcoming a new generation to college: The Millennial students. *Journal of College Admission*, *195*, 20–25. https://eric.ed.gov/?id=EJ783953

Finn, A. N., & Ledbetter, A. M. (2013). Teacher power mediates the effects of technology policies on teacher credibility. *Communication Education*, *62*(1), 26–47. https://doi.org/10.1080/03634523.2012.725132

Finn, A. N., & Ledbetter, A. M. (2014). Teacher verbal aggressiveness and credibility mediate the relationship between teacher technology policies and perceived student learning. *Communication Education*, *63*(3), 210–234. https://doi.org/10.1080/03634523.2014.919009

Fisher, D., Henderson, D., & Fraser, B. (1995). Interpersonal behaviour in senior high school biology classes. *Research in Science Education*, *25*(2), 125–133. https://doi.org/10.1007/BF02356447

Frey, T. K., Moore, K. P., & Dragojevic, M. (2021). Syllabus sanctions: Controlling language and fairness as antecedents to students' psychological reactance and intent to comply. *Communication Studies*, *72*(3), 456–473. https://doi.org/10.1080/10510974.2021.1876130

Frey, T. K., & Tatum, N. T. (2016). Hoverboards and "hovermoms": Helicopter parents and their influence on Millennial students' rapport with instructors. *Communication Education*, *65*(3), 359–361. https://doi.org/10.1080/03634523.2016.1177846

Goodboy, A. K., & Kline, R. B. (2017). Statistical and practical concerns with published communication research featuring structural equation modeling. *Communication Research Reports*, *34*(1), 68–77. https://doi.org/10.1080/08824096.2016.1214121

Goodboy, A. K., & Myers, S. A. (2015). Revisiting instructor misbehaviors: A revised typology and development of a measure. *Communication Education*, *64*(2), 133–153. https://doi.org/10.1080/03634523.2014.978798

Hayes, J. G. (2005). Problematic student behaviors in the college communication classroom: Reviewing and re-envisioning instructional communication research. *Basic Communication Course Annual*, *17*(1), 43–93. http://ecommons.udayton.edu/bcca/vol17/iss1/7

Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods*, *1*(1), 104–121. https://d oi.org/10.1177%2F109442819800100106

Horan, S. M., Martin, M. M., & Weber, K. (2012). Understanding emotional response theory: The role of instructor power and justice messages. *Communication Quarterly*, *60*(2), 210–233. https://doi.org/10.1080/01463373.2012.669323

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185. https://doi.org/10.1007/BF02289447

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

Jenkins, P. H. (1997). School delinquency and the school social bond. *Journal of Research in Crime and Delinquency*, *34*(3), 337–367. https://doi.org/10.1177/0022427897034003003

Jiang, Y., Lee, C. K. J., Wan, Z. H., & Chen, J. (2021). Stricter teacher, more motivated students? Comparing the associations between teacher behaviors and motivational beliefs of western and east Asian learners. *Frontiers in Psychology*, *11*, 3743. https://doi.org/10.3389/fpsyg.2020.564327

Kearney, P., Plax, T. G., Hays, E. R., & Ivey, M. J. (1991). College teacher misbehaviors: What students don't like about what teachers say and do. *Communication Quarterly*, *39*(4), 309–324. https://doi.org/10.1080/01463379109369808

Kendall, K. D., & Schussler, E. E. (2013a). More than words: Probing the terms undergraduate students use to describe their instructors. *International Journal of Teaching and Learning in Higher Education*, *25*(2), 200–212. https://www.isetl.org/ijtlhe/pdf/IJTLHE1528

Kendall, K. D., & Schussler, E. E. (2013b). Evolving impressions: Undergraduate perceptions of graduate teaching assistants and faculty members over a semester. *CBE—Life Sciences Education*, *12*(1), 92–105. https://doi.org/10.1187/cbe.12-07-0110

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). Guilford.

Larseingue, M., Sawyer, C. R., & Finn, A. N. (2012). Components of students' grade expectations for public speaking assignments. *Communication Education*, *61*(4), 428–447. https://doi.org/10.1080/03634523.2012.708765

Ledbetter, A. M. (2009). Measuring online communication attitude: Instrument development and validation. *Communication Monographs*, *76*(4), 463–486. https://doi.org/10.1080/03637750903300262

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.

Lindlof, T. R., & Taylor, B. C. (2017). *Qualitative communication research methods* (4th ed.). Sage.

Martin, M. M., & Anderson, C. M. (1998). The cognitive flexibility scale: Three validity studies. *Communication Reports*, *11*(1), 1–9. https://doi.org/10.1080/08934219809367680

Martin, M. M., Dunleavy, K. N., & Kennedy-Lightsey, C. D. (2010). The instrumental use of verbally aggressive messages. In T. Avtgis & A. S. Rancer (Eds.), *Arguments, aggression, and conflict: New directions in theory and research* (pp. 400–416). Routledge.

Martin, M. M., & Rubin, R. B. (1994). Development of a communication flexibility measure. *Southern Communication Journal*, *59*(2), 171–178. https://doi.org/10.1080/10417949409372934

Martin, M. M., & Rubin, R. B. (1995). A new measure of cognitive flexibility. *Psychological Reports*, *76*(2), 623–626. https://d oi.org/10.2466%2Fpr0.1995.76.2.623

Mazer, J. P. (2012). Development and validation of the student interest and engagement scales. *Communication Methods and Measures*, *6*(2), 99–125. https://doi.org/10.1080/19312458.2012.679244

Mazer, J. P., & Graham, E. E. (2015). Measurement in instructional communication research: A decade in review. *Communication Education*, *64*(2), 208–240. https://doi.org/10.1080/03634523.2014.1002509

Mazer, J. P., Murphy, R. E., & Simonds, C. J. (2007). I'll see you on "Facebook": The effects of computer-mediated teacher self-disclosure on student motivation, affective learning, and classroom climate. *Communication Education*, *56*(1), 1–17. https://doi.org/10.1080/03634520601009710

McCroskey, J. C. (1998). *An introduction to communication in the classroom* (2nd ed.). Tapestry Press.

Moore, K. P., & Richards, A. S. (2019). The effects of instructor credibility, grade incentives, and framing of a technology policy on students' intent to comply and motivation to learn. *Communication Studies*, *70*(4), 394–411. https://doi.org/10.1080/10510974.2019.1617761

Myers, S. A., & Bryant, L. E. (2004). College students' perceptions of how instructors convey credibility. *Qualitative Research Reports in Communication*, *5*, 22–27.

Myers, S. A., & Knox, R. L. (2000). Perceived instructor argumentativeness and verbal aggressiveness and student outcomes. *Communication Research Reports*, *17*(3), 299–309. https://doi.org/10.1080/08824090009388777

Plax, T. G., Kearney, P., McCroskey, J. C., & Richmond, V. P. (1986). Power in the classroom VI: Verbal control strategies, nonverbal immediacy, and affective learning. *Communication Education*, *35*(1), 43–55. https://doi.org/10.1080/03634528609388318

Poplin, M., Rivera, J., Durish, D., Hoff, L., Kawell, S., Pawlak, P., Hinman, I. S., Straus, L., & Veney, C. (2011). She's strict for a good reason: Highly effective teachers in low-performing urban schools. *Phi Delta Kappan*, *92*(5), 39–43. https://doi.org/10.1177/003172171109200509

Radziszewska, B., Richardson, J. L., Dent, C. W., & Flay, B. R. (1996). Parenting style and adolescent depressive symptoms, smoking, and academic achievement: Ethnic, gender, and SES differences. *Journal of Behavioral Medicine*, *19*(3), 289–305. https://doi.org/10.1007/BF01857770

Revelle, W. R. (2017). *Psych: Procedures for personality and psychological research*. Northwestern University. https://www.scholars.northwestern.edu/en/publications/psych-procedures-for-personality-and-psychological-research

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

RStudio Team. (2021). RStudio: Integrated Development Environment for R. RStudio, PBC. http://www.rstudio.com/

Sloat, E., Amrein-Beardsley, A., & Sabo, K. E. (2017). Examining the factor structure underlying the TAP system for teacher and student advancement. *AERA Open*, *3*(4), 1–18. https://doi.org/10.1177%2F2332858417735526

Spitzberg, B. H., & Cupach, W. R. (1984). *Interpersonal communication competence*. Sage.

Spitzberg, B. H., & Cupach, W. R. (1989). *Handbook of interpersonal competence research*. Springer-Verlag.

Steinberg, L. (2001). We know some things: Parent–adolescent relationships in retrospect and prospect. *Journal of Research on Adolescence*, *11*(1), 1–19. https://doi.org/10.1111/1532-7795.00001

Steinberg, L., Blatt-Eisengart, I., & Cauffman, E. (2006). Patterns of competence and adjustment among adolescents from authoritative, authoritarian, indulgent, and neglectful homes: A replication in a sample of serious juvenile offenders. *Journal of Research on Adolescence*, *16*(1), 47–58. https://doi.org/10.1111/j.1532-7795.2006.00119.x

Stiff, J. B., Dillard, J. P., Somera, L., Kim, H., & Sleight, C. (1988). Empathy, communication, and prosocial behavior. *Communication Monographs*, *55*(2), 198–213. https://doi.org/10.1080/03637758809376166

Tatum, N. T., & Frey, T. K. (2021). (In) flexibility during uncertainty? Conceptualizing instructor strictness during a global pandemic. *Communication Education*, *70*(2), 214–216. https://doi.org/10.1080/03634523.2020.1857419

Tatum, N. T., Olson, M. K., & Frey, T. K. (2018). Noncompliance and dissent with cell phone policies: A psychological reactance theoretical perspective. *Communication Education*, *67*(2), 226–244. https://doi.org/10.1080/03634523.2017.1417615

Teven, J. J., & McCroskey, J. C. (1997). The relationship of perceived teacher caring with student learning and teacher evaluation. *Communication Education*, *46*(1), 1–9. https://doi.org/10.1080/03634529709379069

Thomas, C. E., Richmond, V. P., & McCroskey, J. C. (1994). The association between immediacy and socio-communicative style. *Communication Research Reports*, *11*(1), 107–114. https://doi.org/10.1080/08824099409359946

Thompson, B. (2007). The syllabus as a communication document: Constructing and presenting the syllabus. *Communication Education*, *56*(1), 54–71. https://doi.org/10.1080/03634520601011575

Thweatt, K. S., & McCroskey, J. C. (1998). The impact of teacher immediacy and misbehaviors on teacher credibility. *Communication Education*, *47*(4), 348–358. https://doi.org/10.1080/03634529809379141

Tickle, L. (2017, September 18). How strict is too strict at school? *The Guardian*. https://www.theguardian.com/education/2017/sep/18/how-strict-is-too-strick-great-yarmouth-academy-smile

Trefzger, J. W. (2018, August 27). Advice for teachers: Dare to be strict. *Faculty Focus*. https://www.facultyfocus.com/articles/effective-classroom-management/advice-teachers-dare-strict/

Vallade, J. I., & Kaufmann, R. (2018). Investigating instructor misbehaviors in the online classroom. *Communication Education*, *67*(3), 363–381. https://doi.org/10.1080/03634523.2018.1467027

Worley, D., Titsworth, S., Worley, D. W., & Cornett-DeVito, M. (2007). Instructional communication competence: Lessons learned from award-winning teachers. *Communication Studies*, *58*(2), 207–222. https://doi.org/10.1080/10510970701341170

Wubbels, T., Brekelmans, M., & Hooymayers, H. P. (1991). Interpersonal teacher behavior in the classroom. In B. J. Fraser & H. J. Walberg (Eds.), *Educational environments: Evaluation, antecedents and consequences* (pp. 141–160). Pergamon Press.

Wubbels, T., Créton, H., & Hooymayers, H. (1985, March 31–April 4). *Discipline problems of beginning teachers, interactional teacher behavior mapped out*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Wubbels, T., & Levy, J. (1991). A comparison of interpersonal behavior of Dutch and American teachers. *International Journal of Intercultural Relations*, *15*(1), 1–18. https://doi.org/10.1016/0147-1767(91)90070-W

Zhang, Q., & Zhu, W. (2008). Exploring emotion in teaching: Emotional labor, burnout, and satisfaction in Chinese higher education. *Communication Education*, *57*(1), 105–122. https://doi.org/10.1080/03634520701586310