Peer-reviewed version available at *Soc. Sci.* **2018**, *7*, 49; doi:10.3390/socsci7030049

Member Perceptions as a Parsimonious Representation of Collaboration Viability

Benjamin Heslop benjamin.heslop@uon.edu.au[1]
Kylie Bailey kylie.bailey@newcastle.edu.au[1]
Jonathan Paul jonathan.paul@newcastle.edu.au[1]
Elizabeth Stojanovski elizabeth.stojanovski@newcastle.edu.au[2]
[1]School of Medicine and Public Health, Faculty of Health and Medicine
[2]School of Mathematical and Physical Sciences, Faculty of Science, University of Newcastle, Callaghan Campus, *University Drive, Callaghan, NSW 2308, Australia*

Correspondence to Benjamin Heslop
Benjamin's PhD scholarship paid by the School of Medicine and Public Health
Work carried out at Mothers and Babies, Hunter Medical Research Institute

Abstract
The PILAR model provides a dynamical systems perspective on collaboration. Two studies are performed using peer assessment data, both testing empirical support for the five Pillars (*prospects, involved, liked, agency, respect*) that constitute member's perceptions of collaboration viability.
The first study analyses peer assessment data collected online from 458 first-year engineering students (404 males; 54 females). A nine-item instrument was inherited from past year's usage in the course, expanded with four additional items to elaborate upon the *agency* and *liked* Pillars. Exploratory factor analysis was conducted on student responses to test whether they thematically aligned to constructs consistent with the five Pillars. As anticipated, twelve of the thirteen items grouped into five components, each aligned with a Pillar, providing empirical evidence that the five Pillars represent perceptions of collaboration.
The second study replicated the first study using a retrospective analysis of 87 items included in the Comprehensive Assessment of Team Member Effectiveness (CATME) peer assessment tool. The associated factor analyses resulted in five components and conceptual alignment of these components with Pillars was evident for three of five CATME components. We recommend a peer assessment instrument based upon PILAR as potentially more parsimonious and reliable than an extensive list of behaviours, such as employed by CATME. We also recommend including items that target inter-rater bias, which is aligned with the *liked* Pillar, that instruments such as CATME exclude.

Keywords: PILAR, CATME, collaboration, peer assessment, inter-rater bias

Peer-reviewed version available at Soc. Sci. 2018, 7, 49; doi:10.3390/socsci7030049

Introduction

Egalitarian collaboration within hierarchical organisations leads to employee retention and organisational responsiveness (Baker, Grinstein, & Harmancioglu, 2016). Long-term benefits to the organisation have been demonstrated by Appreciative Inquiry, an intervention that creates collaboration outside normal workflow processes, across organisational units, and between levels of hierarchy (MacCoy, 2014). However, unless culturally-embedded and hierarchically-supported, collaboration outside normal workflows may eventually cease, as participants fail to see an impact for their efforts (Edmondson, 2004; Leana & Buren, 1999). Yet, it is not advisable to mandate or incentivise collaboration, since participants may halt once either is removed (James, Mann, & Creasy, 2007; Wageman & Gordon, 2005). To embed viable collaboration within the organisation, we postulate the importance of establishing a definition that serves both participants' bottom-up, and policy designers' top-down, requirements and perspectives. This research aims to attain empirical evidence for a model of collaboration designed to measure collaboration viability. Such a model may also aid design of pro-collaboration organisational policies, based upon how collaboration is perceived by its members.

PILAR Model of Collaboration

Only a generically applicable theory of collaboration can reliably inform wider organisational policy. While numerous theorists have expressed their desire for a universal, generic model of collaboration (Salas, Shuffler, Thayer, Bedwell, & Lazzara, 2015), currently no agreed version exists (Keyton, 2016), and some contend that such a model will never exist (Green, 2015). Designed to be consistent with organisational psychology (Heslop, Paul, Drew, Bailey, & Stojanovski, 2018) and social psychology (Heslop et al., *in press*), the PILAR model of collaboration is a contender for universality.

PILAR considers that humans have an array of perceptions that guide their decision to engage in a group, or to withdraw when membership of the group is no longer in their interest (Wilson, Ostrom, & Cox, 2013). A perspective of tension between the member's pro-social engagement with the group, and self-interested withdrawal, indicates a potential evolutionary basis of collectivism and individualism (Huang & Bargh, 2014). The *team work engagement* that each member feels as a result of assessing their peers; the external environment; and other factors, influences their preparedness to commit personal resources to the team (Costa, Passos, & Bakker, 2014).

The PILAR model asserts that its *Pillars* (*prospects*, *involved*, *liked*, *agency* and *respect*) provide some of the basis upon which the typical collaborating actor is instinctively prompted to remain within, or leave, a team (Oakley & Halligan, 2017). Since these perceptions will influence employee's willingness to collaborate, they should be considered in the organisation's design of any policy or scheme designed to foster collaboration (Table 1).

Table 1 Definition of each Pillar

Pillar	Member's perception of collaboration
prospects	Your opinion of whether the group will succeed, and if so, whether you will receive your anticipated share of that success
involved	Your willingness to cooperate with colleagues, either providing or receiving assistance: in the form of knowledge and physical aid
liked	Your sense of popularity and security based upon your perception of colleagues' warmth and affection toward you
agency	The permission you feel to suggest change to the group's norms, processes, task allocation and strategy
respect	Your opinion of a colleagues' task-relevant competence, and general trustworthiness

Perception of prospects

Peer-reviewed version available at Soc. Sci. 2018, 7, 49; doi:10.3390/socsci7030049

The *prospects* Pillar is the likelihood of the collaboration achieving its goal, and the team member subsequently receiving the anticipated benefit, for instance, produce from a community garden. When the member feels that their group is likely to fail, low *prospects* are experienced as uneasiness and foreboding. Even when the group itself is performing well, if the member suspects their share is at risk, then that member perceives *prospects* as low (Heslop, Bailey, Paul, Drew, & Smith, 2016).

Perception of involved

The *involved* Pillar encourages two members to cooperate in providing knowledge or physical assistance to complete a task. It is experienced by the member as an openness to, and comfort working with, a specific colleague, which at high levels is experienced as enthusiasm to participate (Quinn & Dutton, 2005). Lack of the *involved* perception is experienced as general trepidation and unease, due to the potential risk of cooperation, such as embarrassment at needing help, an unappreciative recipient of advice, or their becoming a competitor for expert status (Klein, Lim, Saltz, & Mayer, 2004).

Perception of liked

The *liked* Pillar is associated with belonging and security, whereas being disliked leads to feelings of social isolation and insecurity. Those who perceive they are poorly *liked* may reduce the unpleasant feeling by mending relationships, or if this is not practical or desired, leaving the group entirely. This behaviour is prominent for those who hold an identity linked to their group, who therefore cannot tolerate being disliked by in-group colleagues, and willingly engage in out-group exclusion (Meeussen, Delvaux, & Phalet, 2014).

Perception of agency

The *agency* Pillar describes feeling empowered to suggest change to the collaboration, for example, a strategy change based upon foreseen dangers (Spânu et al., 2013), or suggesting that task responsibilities between members be adjusted.

Perception of respect

The *respect* Pillar reflects the member's perception of a colleague's competence and character (Ibrahim & Ribbers, 2009). High *respect* is faith and trust in colleagues' dependability, compared to low *respect*, associated with distrust and vigilance (Ko, 2010).

Hypotheses Related to Peer Assessment Data

Peer assessment is a survey of employees' opinions of their colleagues, commonly implemented by organisations as an adjunct to supervisor-assessment. Unlike PILAR, peer assessment data is not intended to evaluate the collaboration, but the data may nevertheless exhibit themes reminiscent of the five Pillars. The presence of themes within peer assessment that align with the five Pillars may therefore constitute empirical support for PILAR. Using peer assessment data, both of the studies in this article will test both of the following hypotheses.

Hypothesis one

While the *respect* and *involved* Pillars result from cumulative perceptions of individual peers, the remaining Pillars may be highly affected by some individuals, yet not at all by others. *Liked*, *agency* and *prospects* Pillars may not be affected by every peer, but rather, only by those whose behaviour makes an inordinate impact on the member, whether positive and negative. For instance, the member's perception of *agency* may be damaged by a single overbearing colleague, or their perception of *liked* may be increased due to a single affirmative comment.

Nevertheless, we postulate that all Pillars will be evident in data derived from peers' assessment of one another, due to the impact upon whether the peer's actions positively or negatively affect the respondent's willingness to engage with the group. This is because, as previously noted, each member's engagement is based upon their perception of all Pillars, with all peers potentially contributing to each. For instance, group disengagement, otherwise resulting from a single rude peer (liked), might be counter balanced by another colleague's

Peer-reviewed version available at Soc. Sci. 2018, 7, 49; doi:10.3390/socsci7030049

(or even the same person's) trustworthiness (*respect*). Thus, each member's perception of each Pillar may be influenced by the behaviour and attributes of any peer. This leads to the first hypothesis that respondents will give each peer a similar assessment to items that thematically align with each Pillar.

Hypothesis two

A further implication of PILAR is that the member's decision to increase or reduce engagement in the collaboration depends upon the relative importance they place on each of the five Pillars. Relative weighting will vary depending upon the collaboration context, and the personalities prevalent within the cohort. (Salas, 2005). Weighting may vary according to the situation, for example, a military team focussing on *prospects* of surviving, compared to a gardening club wishing to maintain relationships (*liked*). Weighting of Pillars may also vary according to the member's personality and psychological resources (Luthans, Youssef, & Avolio, 2007), for instance, a confident member is less concerned with being *liked*, and those with less optimism in others' intentions might be less interested in demonstrating *agency* (Heslop et al., *in press*). This leads to the second hypothesis that the order of exploratory components, from most to least explanatory, will reflect the priority that the cohort assign to each component.

Materials and Methods

In the first study, both hypotheses were tested using an exploratory factor analysis of engineering student responses to the peer assessment instrument. To test hypothesis one, we consider predictions based upon hypothesised alignment between peer assessment data and Pillars. To test hypothesis two, we determine which components will explain the most variance in student response data, based upon conceptual alignment between Pillars and engineering students' personality profile.

Participants

In teams of three to six, engineering students from an Australian university participated in a design project as part of their 2016 first semester's coursework. At the end of semester, each student completed a peer assessment instrument for each of their team mates, using the SPARK^{PLUS} online platform (Willey & Gardner, 2010), which is henceforth referred to as the PILAR instrument.

Instructions to Survey Participants

Before attempting the survey, students received written instructions on how to use the survey interface and how the group's data might be used to adjust their mark. Students were cautioned by the lecturer that overestimating their own contribution at the expense of others may result in their evaluations being omitted from the analysis. Students were also advised that the total pool of marks available to the group members was fixed, and if they chose to rate another member highly, they were therefore in effect giving permission to the lecturer to take some of their own marks and award them to the high-performing member of the group. Students were also asked to mark 'average' behaviour at 70 out of a possible 100, and to use very low or very high marks for exceptionally poor or favourable teamwork behaviours, respectively.

PILAR peer assessment instrument

We inherited a nine-item peer assessment instrument from the previous year's application of it by the lecturer. While the original instrument contained items related to peers' competence (*respect*), motivation (*prospects*) and teamwork (*involved* and *liked*), at the authors' request, the instrument was expanded to include four additional items that aligned with *agency*, *involved* and *liked* (Table 2).

Table 2 Peer assessment questions administered by the online PILAR instrument. For each question, alignment with Pillars, and locus of perception, is nominated.

Q	PILAR Instrument Item	Pillar	Locus of perception

Peer-reviewed version available at Soc. Sci. 2018, 7, 49; doi:10.3390/socsci7030049

Q1	How much work and effort has the person put into the project? Have they done their fair share and pulled their weight?	*prospects*	respondent
Q2	How well did the person get their work completed by the agreed upon time?	*prospects*	respondent
Q3	Did the person regularly attend team meetings on time?	control	
Q4	How good is the quality of the person\'s research and their understanding of what they have read?	*respect*	respondent
Q5	How good are the person\'s analytical and problem-solving skills?	*respect*	respondent
Q6	How good are the person\'s reporting writing and editing skills?	*respect*	respondent
Q7	Did you find this person easy to interact with?	*liked*	respondent
Q8	How well did the person contribute to discussion during team meetings?	*agency*	peer
Q9	How well did the person encourage others and value their ideas?	*involved*	peer
Q10	How well did the person assist others when asked?	*involved*	peer
Q11	Did this person express their opinions with confidence?	*agency*	peer
Q12	Did the person constructively challenge other people's opinions?	*involved*	peer
Q13	Was this person well-liked by the group?	*liked*	peer

Although inheritance of the first nine, unmodified, items prevented the authors from specifically designing questions to align with Pillars, data were nonetheless assessed for evidence of consistency with Pillar perceptions within this cohort. As it does not necessarily indicate a peer or respondent's perception of the group, Q3 as a measure of punctuality was not considered to align logically with a Pillar. Apart from commitment to the project, punctuality is influenced by a range of factors, such as whether the peer is organised, cultural background, and schedule conflicts (Back, Schmukle, & Egloff, 2006; Basu & Weibull, 2002). Q3 was therefore considered a control question, and the factor analysis was expanded to six components to accommodate it. The purpose of Q3 as a control is to demonstrate that the components derived from the factor analysis align with the Pillars, even with the inclusion of an item that does not conceptually align to a Pillar.

Removal of Outliers

As part of standard instructions given by the instrument, for each item students were instructed to rate the average collaborator as 70/100. We therefore contend that a rating five or less out of 100 may be influenced by *inter-rater bias* (Magin, 2001; Viswesvaran, Schmidt, & Ones, 2005). Inter-rater bias occurs where a personal relationship between peer and respondent bias peer assessments ratings made by the respondent, leading to rating on a primarily subjective basis, rather than on the peer's objective behaviour. For this reason, we remove from analysis any set of 13 ratings that includes any rating of five or less. We consider it highly unlikely that such a low rating is an objective assessment, and is rather prompted by the respondent's antipathy for the peer.

Hypothesis predictions

Predicted alignment of PILAR instrument with Pillars

Before aligning PILAR items into Pillars, we note that the perception being rated may be that of the peer, or the respondent. Q1 and Q2 align with the respondent's perception of group *prospects*, which are reduced by unmotivated and free-riding peers. Q4, 5 and 6 align with the respondent's perception of *respect*, because all three competencies relate to the respondent's view of the peer's abilities. Q7 asks if the respondent feels liked by the peer,

Peer-reviewed version available at Soc. Sci. 2018, 7, 49; doi:10.3390/socsci7030049

based upon how comfortable their interaction was. Q13 considers that whether a peer feels disliked, which may be inversely related to whether the respondent feel disliked by that peer. Q8 and Q11 align with the peer's perception of *agency*, and whether they felt able to express an opinion or contribute to the discussion. Q9, Q10 and Q12 align with the peer's perception of *involved,* in terms of whether they are willing help, or be helped by, colleagues. For example, Q9 asks whether a peer asks for a colleague's opinion, which necessitates the peer's willingness to accept assistance.

Predicted component priority due to cohort personality

In their survey of 4,876 engineers and 75,892 non-engineers, Williamson, Lounsbury, & Han (2013) discovered that engineers' personality could be classified as more tough-minded (assess problems logically) and intrinsically-motivated (seeking challenge, meaning and autonomy), but lacking assertiveness (ability to speak up), service orientation (meeting collaborator's needs), emotional stability (resilience), extraversion (sociability), optimism (positive outlook), image management (control one's presentation to others), visionary style (forward thinking) and work drive (dedication to project goals).

Considering these attributes of engineers, we posit that *respect* will be of primary importance because of it requires logical assessment of colleagues' competence, whereas remaining assessment items are more qualitative. We consider that *prospects* will also be important, since the engineer's intrinsic motivation will produce a preference for collaboration in which colleagues are also highly dedicated. It is difficult to be autonomous, find meaning and be challenged when colleagues require oversight because of their poor skills and motivation.

Conversely, lack of assertiveness, visionary style and optimism indicates that *agency* will be a lower-priority perception, and therefore the component less relevant. Taking optimism as an example, without believing that change is possible, there is no point suggesting it (Heslop, Stojanovski, Paul, Bailey, & Drew, 2018). A desire to be *involved* would require a service orientation and work drive; the latter because aiding colleague's places project goals above one's own interests. Finally, we consider that *liked* would be important to those with extraversion and concerned with image management. Hence, within this cohort, we expect that *respect* and *prospects* will feature among the first two components, with *involved*, *agency* and *liked* among the latter three.

Results

In teams of between three and six, 458 first-year engineering students completed the PILAR peer assessment instrument (404 males; 54 females, mean age 20.1 years, S.D. 4.6 years, and 91.9 % Australian national). Descriptive statistics and correlations between the 13 items are shown in Table 3. Each item's mean is over 80, and the distribution of each item is negatively skewed. As expected, components were highly correlated (r = 0.67 – 0.80; the least being component six; r = 0.54 – 0.65) due to the positive-feedback contended to exist between Pillars, noted previously.

Table 3 Mean and Standard Deviation of each item's ratings, and Pearson correlations between items. All correlations are significant ($p < 0.01$) and each item has 2228 scores.

Descriptive			**Pearson Correlation Coefficients**											
Q	Mean	S.D.	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12
Q1	82.9	11.9	1											
Q2	82.9	12.2	.83	1										
Q3	85.8	11.4	.64	.62	1									
Q4	83.3	11.6	.83	.74	.60	1								
Q5	83.1	10.6	.78	.74	.60	.84	1							
Q6	82.6	11.8	.79	.71	.56	.82	.79	1						
Q7	85.9	11.3	.69	.63	.55	.70	.69	.67	1					
Q8	84.0	11.6	.76	.69	.59	.77	.79	.74	.77	1				
Q9	83.8	10.8	.68	.61	.52	.68	.73	.66	.76	.76	1			

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 February 2018 doi:10.20944/preprints201802.0183.v1

Peer-reviewed version available at Soc. Sci. 2018, 7, 49; doi:10.3390/socsci7030049

Descriptive			Pearson Correlation Coefficients											
Q10	84.1	10.8	.76	.70	.60	.77	.76	.74	.73	.78	.77	1		
Q11	84.5	10.4	.68	.64	.56	.71	.76	.67	.68	.80	.73	.75	1	
Q12	81.9	11.4	.69	.62	.52	.72	.76	.70	.65	.75	.75	.74	.76	1
Q13	85.6	11.5	.72	.67	.59	.72	.73	.69	.83	.77	.75	.76	.71	.68

Factor analysis was permissible as the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was above the recommended cut-off of 0.6 (KMO = 0.97) and Bartlett's test of Sphericity was statistically significant ($p < 0.05$). The factor analysis was set at six, based upon five Pillars plus a control component. Kaiser Normalisation was applied, using Promax rotation due to the high correlation coefficients ($r > 0.6$) between items (Dean, 2009). Promax was also chosen because Pillars are theorised to be correlated as a result of positive reinforcing feedback between four pairs (*prospects* and *involved*; *involved* and *respect*; *liked* and *involved*; and *respect* and *prospects*), with stabilising feedback between the remaining six pairs of Pillars (Heslop et al., *in press*).

The factor analysis converged in six iterations and the resulting factors cumulatively explained 90 % of the variance of student responses. Each of the resulting factors had strong scale reliability (Cronbach's alpha > 0.8) indicating that items making up each factor were closely related. Only items with factor loadings over 0.5 were considered for inclusion in each scale.

Eleven of the twelve possible PILAR items aligned to one of the five components as predicted by PILAR (Table 4). The single exception was Q10's loading of 0.40 (below the 0.5 cut-off) in component three. Nevertheless, Q10 still has its largest loading aligned with involved. The non-Pillar control item regarding punctuality, Q3, dominated the sixth component. We therefore conclude that the first study supports hypothesis one, and that Pillars are a parsimonious classification of perceptions of collaboration.

Table 4 Pattern matrix (loadings > 0.5) resulting from exploratory factor analysis (six components). Cronbach's Alpha and percentage of variance is included. The control question (Q3) represents a sixth component.

	Component loadings					
	1 *respect*	2 *liked*	3 *involved*	4 *prospects*	5 *agency*	6 Control (Q3)
Q1				.58		
Q2				1.03		
Q3						1.00
Q4	.86					
Q5	.63					
Q6	1.10					
Q7		.96				
Q8					.54	
Q9			.84			
Q10			(.40)			
Q11					.96	
Q12			.75			
Q13		.79				
Alpha	.93	.91	.85	.90	.89	
Variance (%)	73.24	5.24	4.06	3.34	2.32	2.10

Discussion

Factor analysis of the PILAR data partially supported hypothesis one that peer assessment data can be parsimoniously captured through the PILAR perceptions of collaboration. However, with the first component explaining almost three quarters of the variance, later components are susceptible to small permutations in the data. Therefore,

further work is required to develop an instrument that might generate data with more variance-balanced, but less correlated, components that can claim to represent the five Pillars. However, the extent of balance achievable may be limited by inter-rater bias, which Viswesvaran et al. (2005) estimated to explain 60% of peer assessment rating variance.

Hypothesis two concerned component order, and whether it suggests teamwork priorities of the cohort. Capturing more variance may indicate that respondents placed greater importance on answering the component's constituent items accurately. The first two components were *respect* followed by *liked,* which only partly supports Hypothesis two; that predicted *respect* followed by *prospects*.

One potential explanation for the *liked* component being second priority is that inter-rater bias was captured by items aligned with *liked*. This may have been exacerbated by interpersonal conflict arising from engineers' lack of emotional stability (resilience), extraversion (sociability) and image management (control one's presentation to others) (Williamson et al., 2013).

The fourth place for the *prospects* component may have been a result of this cohort not being representative of professional engineers. The attrition rate of first year undergraduate engineers is typically over 50 % (Felder, Felder, & Dietz, 1998), implying that half of this cohort will choose a different career. Due to the Dunning-Krueger effect, this may have led to unusual interpretations for items aligned with *prospects*, since less-competent members evaluate peer's performance less accurately (Schlösser, Dunning, Johnson, & Kruger, 2013). Since *prospects*-aligned questions in the instrument are more subjective than *respect*-aligned questions, they may have suffered from greater inaccuracy from those students destined to not achieve an undergraduate engineering degree.

Study Limitations

There are numerous limitations to our approach that might be remedied in future. First, to reduce signalling that may have artificially induced respondents to answer in alignment with Pillars, the first six items would ideally be mixed among the remaining items rather than sequenced (Fowler, 1995). Second, to reduce respondents' superficially giving identical answers to similarly-phrased items, items would ideally be phrased differently, for example *respect*-aligned. Third, a future replication of this research would improve question design (Fowler, 1995). For instance, Q1 and Q2 both have multiple concepts in one item. A separate investigation of the PILAR instrument found that eight items were poorly designed, and potentially led responds to disengage, hence reduce its reliability (Heslop, Stojanovski, Paul, Iveson, & Bailey, 2018). This may have contributed to the emphatic capturing of variance within the first component.

Another limitation was mixing peer- and respondent-based perceptions of collaboration within Pillars of the first study instrument. A future study might consider focussing solely upon one locus of perception. For instance, in other work, we have proposed an instrument exclusively based upon peer's perception, designed to measure collaboration viability. We consider that this instrument may also be used for peer assessment purposes (Heslop, Stojanovski, Paul, & Bailey, 2017). To address limitations of the first study, we now compare its findings with those of a highly-cited (114 citations) study that employs a similar, factor analysis-based, method.

Second Study

This study considers the initial validation of a free, well-established and widely-used online peer assessment instrument; the Comprehensive Assessment of Team Member Effectiveness (CATME), designed for group projects in educational settings (Ferguson, Lally, Somnooma, Murch, & Ohland, 2016). Re-examining hypothesis one, we consider whether CATME components produced by factor analysis conceptually align with the five Pillars. To investigate such an alignment, we first make post-hoc 'predictions' by conceptually

associating each CATME item either with a: Pillar perception of collaboration; or otherwise a non-Pillar, control construct.

CATME has been selected for this study because it suffers less from limitations present in the PILAR instrument. CATME questions are well-designed as they are drawn from previously-validated instruments. Additionally, we observe that CATME's items are similarly worded. However, we are unsure if aligned CATME items appear randomly or sequentially, since the original instrument given to respondents is not published.

Method of Loughry et al (2007)

Loughry, Ohland, & Moore (2007) developed CATME by first formulating a list of 392 items, drawn from pre-existing instruments, for example measuring *team potency* and *Team Climate Inventory* (Anderson & West, 1998). Fifteen colleagues of Loughry et al. (2007) then selectively reduced these 392 to a shorter list of 218 by removing poorly formed and unclear items. Eighty-seven of the most significant were then identified by surveying 2,777 students, who were asked to rate the importance of each peer behaviour to team success. Finally, and similar to the first study, 1,157 students were each asked to rate a collaborator they remembered on each of the 87 items (listed in Loughry et al., 2007).

The variance of this data was tested first via an exploratory-, followed by confirmatory-, factor analysis with a one-, five- and seven-component model. Five components best explained the variance, however they exhibited high inter-correlation (r = 0.77 – 0.93) that the authors claimed required further investigation. While Loughry et al. (2007) found that five components were superior to one and seven, it is possible that the data could be better explained by fewer components than five; by six components; or by more than seven.

Predicted alignment of CATME items with Pillars

Our examination of the 87 CATME items suggests that 66 items align with four Pillars. The remaining 21 items align with leadership, which we considered to be a control. No CATME items align with *liked*, and only three align with *agency*: "Made important contributions to the team's final product," "Provided insights and ideas that improved the team project," and "Made recommendations that improved the team's performance." Given this, we predict that three CATME components will separately align with *involved*, *respect* and *prospects*, and a control component will align with leadership. We cannot predict how the fifth component might be constituted, given the sparcity of *agency*-aligned items, and lack of *liked*-aligned items.

As expected, items contained within two smaller CATME components clearly align with *respect* and *prospects,* while the largest component with 30 items aligns with *involved* (Table 5). One of the remaining two, larger, components; that with 24 items aligns with all three Pillars just mentioned, and additionally *agency*. The medium-sized component, with 21 items, aligns with leadership, which we consider to be control items that should enter its own component.

Table 5 Five CATME components and their alignment of constituent questions with Pillars. Also, the locus of perception of the CATME component (but not necessarily of all constituent questions).

CATME Component	No. items	Pillar	Locus
Contributing to the team's work	24	*agency, respect, involved, prospects*	Peer
Interacting with teammates	30	*involved*	Peer
Keeping the team on track	21	leadership	
Expecting quality	6	*prospects*	Peer
Having relevant knowledge, skills and abilities	9	*respect*	Respondent

Inter-rater bias causing intercorrelation

One limitation of study two is sparsity of items aligned with *liked* and *agency*, which may also explain high correlations between CATME components. Following on from the first study, with *liking* reaching second-place, and the first component explaining over 70% of the variance, we consider that these correlations might have been a result of a lack of *liked* items. The PILAR instrument had two items related to *liked*, which were included in the second component, potentially indicating its importance.

No CATME items refer to the relationship between peer and respondent, whether from their own, or the peer's perception. Lack of *liked*-aligned items may have encouraged respondents to distribute inter-rater bias felt towards the peer amongst the remaining items, rather than within *liked*-aligned items. This uses the rationale of Freud's *emotional displacement* (Fenichel, 1946), since inclusion of items specifically targetting inter-rater bias may otherwise reduce its effect on the rest of the items, because respondents have a safe, approved channel for their negative or positive affect for the peer.

Relative to the first study, inter-rater bias in the second study may have been exacerbated by numerous factors. The respondent is rating a historical rather than a current colleague, and details of their behaviour may have become uncertain. Further, CATME's 87 items is a long instrument given the wide range of item phrasing adopted, which prompted its later reduction to 33 items (Loughry et al., 2007). Whether from exhaustion from a long instrument, or uncertain recall, it is understanding that inter-rater bias would increase correlation between components.

Also potentially reducing inter-rater bias, the first study used 13 items, repeated for multiple peers, whereas CATME had 87 novel items, which may have increased their cognitive load, and increased the influence of affect in rating peers (Morewedge & Kahneman, 2010). Finally, the first study collected information on current rather than historical peers, presumably allowing sharper details to inform ratings. A shorter instrument, applied to current peers, and with some items aligned to *liked*, may have reduced intercorrelation between components in the first study, otherwise higher in Loughry et al. (2007), perhaps due to inter-rater bias.

Lack of agency and liked in the CATME instrument

Rather than reflecting an inherent lack of importance of *agency* and *liked*, we consider it an artefact of the author's process for deriving the 87 items. *Agency* is only comparatively recent construct, having been initially proposed as *participative safety*, before being popularised by *psychological safety* (Edmondson, 1999). Therefore, *agency* may not have been strongly represented by the instruments from which the 392 items were derived. This sparcity may have been compounded by an introspective, conformist student cohort ascribing less value to peers' *agency* in rating the 218 items to arrive at 87 (Arnett, 2006). The CATME cohort were equal parts engineering and business students, neither of which are noted for empathy that otherwise allows appreciation of others' *agency* (Levenson, Kiehl, & Fitzpatrick, 1995)

In contrast to *agency*, *liked* is a relatively-longstanding construct, originally in the form of *interpersonal attraction*; an aspect of *social cohesion* (Back, 1951), and in the negative sense, *interpersonal conflict* (Exline & Ziller, 1959). A lack of *liked*-aligned items may have been the result of prioritising behaviours rather than perceptions. Of the five Pillars, we consider that *liked* is the most subjective and intangible, followed by *agency*. While the original list of 218 items is not published, there is also selection bias of original selection of items to consider. For instance, 30 items were extracted from Anderson & West's (1998) 60-item Team Climate Inventory. It mapped on to five components, two of which respectively align to *agency* ("participative safety") and *liked* ("interaction frequency"), both of which have over ten constituent items.

There is no published list of the 392 items originally chosen by Loughry et al. (2007), and no published list the subsequent list of 218 derived by 15 colleagues. Hence, we cannot

know at what point *agency*- or *liked*-aligned items were excluded. Nevertheless, their lack obviously prevented a CATME component being aligned with *agency* and *liked*. More broadly, it and inter-rater bias, may explain why there was a single CATME component that encapsulated a range of Pillars, and why the inter-correlation between CATME components was high.

CATME versus PILAR

We note that CATME focusses on peer behaviour, rather than perceptions of the group that influence peer behaviour, yet this approach to peer assessment may offer advantages. First, since perceptions are less diverse than behaviours, a shorter instrument is possible. Second, perceptions are not limited by natural ability, whereas behaviours are, which the respondent must evaluate. Evaluating behaviours may be more prone to inter-rater bias than evaluating peer's perceptions, since the latter require an empathetic rather than judgemental perspective (Oakley & Halligan, 2017). Finally, since behaviours may be difficult to calibrate between respondents, to improve accuracy of a survey, training to use a behaviour-based peer assessment instrument is recommended (van Zundert, Sluijsmans, & van Merrienboer, 2010). By contrast, a perception-based instrument, if based upon familiar, universal perceptions, may not require training for respondents before use (Heslop et al., 2017).

While behaviour-based peer assessment cannot ignore personality variation; a perception-based instrument founded upon universal perceptions must do so. In other words, for PILAR, personality is a confounder, whereas for CATME, perceptions of collaboration become confounders. For instance, the item: "Provided insights and ideas that improved the team project" might be due to a peer's lack of imagination, or from the Pillar perspective, not feeling they had *agency*. The item "Kept trying when faced with difficult situations" if answered in the negative, might have arisen from a peer perceiving that the group's *prospects* are sufficiently poor to not warrant their continued investment (Graham & Sloan, 2016). Many CATME items can be interpreted as either the respondent's valid judgement of the peer's behaviour, or alternatively the peer's behaviour being justified based upon their perception of the group.

PILAR has taken a different approach to creating an instrument, by encapsulating social psychology theory directly, rather than incorporating instruments currently in use. Since both approaches provide five constructs, we contend that inducing their nature directly from empirically-validated theory is potentially more robust than utilising existing instruments. Not only because of changing trends in the field (Green, 2015), but also due to potential biases within CATME's original creation (Simmons, Nelson, & Simonsohn, 2013). We therefore contend that PILAR may offer a more parsimonious basis upon which to survey respondents' assessment of peers have of the team (Heslop et al., 2017). Indeed, if perceptions are causal to behaviour, it may be that CATME is largely measuring peer's perceptions.

Conclusion

This article presents empirical evidence for the PILAR (*prospects, involved, liked, agency, respect*) model of collaboration. Over 400 engineering students participated in peer assessment as part of their course requirements. Assessment of their responses using factor analysis revealed almost-perfect alignment of the first five components with predicted Pillar-aligned items. The sequence of components may indicate that the cohort prioritise *respect*, followed by *liked*. The latter being unexpected, we considered that either the cohort was not representative of engineers, or that the priority given to the *liked*-aligned component was due to capturing inter-rater bias.

The second study only partially supports PILAR by demonstrating that three Pillars align with three components of CATME's more extensively-developed instrument. Lack of CATME items aligned with *liked* and *agency,* and only examining one, five and seven components, may have precluded potentially stronger endorsement of PILAR. We postulated that methodological biases, and historical trends in literature, may have limited the

numerousness of items aligned with to *agency* and *liked*, leading to high intercorrelation between CATME components.

Despite respective limitations of each study, our findings encourage further exploration of perceptions of collaboration, whether peer's or respondent's, as a method of assessing collaboration viability. Should future studies demonstrate alignment with Pillars, it will constitute further theoretical support for PILAR as a parsimonious model of collaboration. Such a universal model may not only enable organisations to measure the viability of their constituent collaborations with a brief instrument, but also provide guidance for developing policy that fosters egalitarian, voluntary and intrinsically-motivated collaboration; and thereby innovation.

Author Disclosures and Statements

Anonymised data was provided by Dr Simon Iveson from the University of Newcastle in accordance with Australian national standards of research ethics (found at www.oaic.gov.au). Dr Iveson's original peer assessment data was collected in accordance with the 2013 Declaration of Helsinki.

The raw data is available upon request to the corresponding author.

No authors have a conflict of interest or funding sources to report

References

Anderson, N. R., & West, M. A. (1998). Measuring climate for work group innovation: Development and validation of the team climate inventory. *Journal of Organizational Behavior*, *19*(3), 235.

Arnett, J. J. (2006). Suffering, Selfish, Slackers? Myths and Reality About Emerging Adults. *Journal of Youth and Adolescence*, *36*(1), 23–29. https://doi.org/10.1007/s10964-006-9157-z

Back, K. W. (1951). Influence through social communication. *The Journal of Abnormal and Social Psychology*, *46*(1), 9–23. https://doi.org/10.1037/h0058629

Back, M. D., Schmukle, S. C., & Egloff, B. (2006). Who is late and who is early? Big Five personality factors and punctuality in attending psychological experiments. *Journal of Research in Personality*, *40*(5), 841–848. https://doi.org/10.1016/j.jrp.2005.11.003

Baker, W. E., Grinstein, A., & Harmancioglu, N. (2016). Whose Innovation Performance Benefits More from External Networks: Entrepreneurial or Conservative Firms. *Journal of Product Innovation Management*, *33*(1), 104–120. https://doi.org/10.1111/jpim.12263

Basu, K., & Weibull, J. W. (2002). Punctuality - A Cultural Trait as Equilibrium. *Economics for an Imperfect World: Essays in Honor of Joseph E. Stiglitz*, (582), 163–182. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=317621

Costa, P. L., Passos, A. M., & Bakker, A. B. (2014). Team work engagement: A model of emergence. *Journal of Occupational and Organizational Psychology*, *87*, 414–436. https://doi.org/10.1111/joop.12057

Dean, J. (2009). Choosing the Right Type of Rotation in PCA and EFA. *Shiken: JALT Testing & Evaluation SIG Newsletter*, *13*(November), 20–25.

Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, *44*(2), 350–383. Retrieved from http://www.jstor.org/stable/2666999

Edmondson, A. C. (2004). Psychological safety, trust, and learning in organizations: a group-level lens. In R. Kramer & K. Cook (Eds.), *Trust and Distrust in Organizations: Dilemmas and Approaches* (pp. 239–272). New York: Russell Sage Foundation.

Exline, R. V., & Ziller, R. C. (1959). Status Congruency and Interpersonal Conflict in Decision-Making Groups. *Human Relations*, *12*(2), 147–162. https://doi.org/10.1177/001872670053100 4

Felder, R. M., Felder, G. N., & Dietz, E. J. (1998). A Longitudinal Study of Engineering Student Performance and Retention V. Comparison with Traditionally Taught Students. *Journal of Engineering Education*, (October), 469–480. https://doi.org/10.1109/FIE.1992.683370

Fenichel, O. (1946). *The Psychoanalytic Theory of Neurosis*. London: W. W. Norton and Company.

Ferguson, D. M., Lally, C., Somnooma, H. I., Murch, O., & Ohland, M. W. (2016). Using frame-of-reference training to improve the dispersion of peer ratings in teams. *Proceedings - Frontiers in Education Conference, FIE*, *2016–Novem*, 0–5. https://doi.org/10.1109/FIE.2016.7757706

Fowler, F. (1995). *Improving survey questions: Design and evaluation*. Thousand Oaks, California: Sage.

Graham, A., & Sloan, K. (2016). *Resilience at Work: Negotiating a Dynamic World*. *Sloan Group White Papers* (Vol. 8). https://doi.org/10.1136/bmj.322.7301.1536

Green, C. D. (2015). Why Psychology Isn't Unified, and Probably Never Will Be. *Review of General Psychology*, *19*(3), 207–214. https://doi.org/10.1037/gpr0000051

Heslop, B., Bailey, K., Paul, J., Drew, A., & Smith, R. (2016). Collaboration Guidelines to Transform Culture. *Interdisciplinary Journal of Partnership Studies*, *3*(3).

Heslop, B., Paul, J., Drew, A., Bailey, K., & Stojanovski, E. (2018). Organisational Psychology versus Appreciative Inquiry: Unifying the Empirical and the Mystical. *AI Practitioner*, *20*(1), 69–90. https://doi.org/10.12781/978-1-90754934211

Heslop, B., Stojanovski, E., Paul, J., & Bailey, K. (2017). Are We Collaborating Yet? Employee Assessment of Peer's Perceptions. *International Journal of Human Resource Studies*, *7*(4), 175–192.

Heslop, B., Stojanovski, E., Paul, J., & Bailey, K. (2018). PILAR: A Model of Collaboration to Encapsulate Social Psychology. *Review of General Psychology*, *In press*. Retrieved from http://dx.doi.org/10.1037/gpr0000141

Heslop, B., Stojanovski, E., Paul, J., Bailey, K., & Drew, A. (2018). Wellbeing as Cause and Effect of Collaboration. *Manuscript Submitted for Publication*.

Heslop, B., Stojanovski, E., Paul, J., Iveson, S., & Bailey, K. (2018). Respondent disengagement from a peer assessment instrument measuring Collaboration Viability. *Australasian Journal of Engineering Education*, *In press*. https://doi.org/10.1080/22054952.2018.1437676

Huang, J. Y., & Bargh, J. A. (2014). The Selfish Goal: Autonomously operating motivational structures as the proximate cause of human judgment and behavior. *Behavioral and Brain Sciences*, *38*(1), 121–135. https://doi.org/10.1017/S0140525X13000290

James, K. T., Mann, J., & Creasy, J. (2007). Leaders as Lead Learners: A Case Example of Facilitating Collaborative Leadership Learning for School Leaders. *Management Learning*, *38*, 79–94. https://doi.org/10.1177/1350507607073026

Keyton, J. (2016). The Future of Small Group Research. *Small Group Research*, *47*(2), 134–154. https://doi.org/10.1177/1046496416629276

Leana, C., & Buren, H. Van. (1999). Organizational social capital and employment practices. *Academy of Management Review*, *24*(3), 538–555. Retrieved from http://amr.aom.org/content/24/3/538.short

Levenson, M. R., Kiehl, K. A., & Fitzpatrick, C. M. (1995). Assessing psychopathic attributes in a noninstitutionalized population. *Journal of Personality and Social Psychology*, *68*(1), 151–158. https://doi.org/10.1037/0022-3514.68.1.151

Loughry, M. L., Ohland, M. W., & Moore, D. D. (2007). Development of a Theory-Based Assessment of Team Member Effectiveness. *Educational and Psychological Measurement*, *67*(3), 505–524. https://doi.org/10.1177/0013164406292085

Luthans, F., Youssef, C., & Avolio, B. (2007). *Psychological capital: Developing the human competitive edge*. Oxford University Press.

MacCoy, D. J. (2014). Appreciative Inquiry and Evaluation – Getting to What Works. *The Canadian Journal of Program Evaluation*, *29*(2), 104. https://doi.org/10.3138/cjpe.29.2.104

Magin, D. (2001). Reciprocity as a Source of Bias in Multiple Peer Assessment of Group Work. *Studies in Higher Education*, *26*(1), 53–63. https://doi.org/10.1080/0307507002003071

Meeussen, L., Delvaux, E., & Phalet, K. (2014). Becoming a group: Value convergence and emergent work group identities. *British Journal of Social Psychology*, *53*(2), 235–248. https://doi.org/10.1111/bjso.12021

Morewedge, C. K., & Kahneman, D. (2010). Associative processes in intuitive judgment. *Trends in Cognitive Sciences*, *14*(10), 435–440. https://doi.org/10.1016/j.tics.2010.07.004

Oakley, D. A., & Halligan, P. W. (2017). Chasing the Rainbow: The Non-conscious Nature of Being. *Frontiers in Psychology*, *8*(November), 1–16. https://doi.org/10.3389/fpsyg.2017.01924

Quinn, R. W., & Dutton, J. E. (2005). Coordination as Energy-in-Conversation. *The Academy of Management Review*, *30*(1), 36–57.

Salas, E. (2005). Is there a "Big Five" in Teamwork? *Small Group Research*, *36*(5), 555–599. https://doi.org/10.1177/1046496405277134

Salas, E., Shuffler, M. L., Thayer, A. L., Bedwell, W. L., & Lazzara, E. H. (2015). Understanding and Improving Teamwork in Organizations: A Scientifically Based Practical Guide. *Human Resource Management*, *54*(4), 599–622. https://doi.org/10.1002/hrm.21628

Schlösser, T., Dunning, D., Johnson, K. L., & Kruger, J. (2013). How unaware are the unskilled? Empirical tests of the "signal extraction" counterexplanation for the Dunning-Kruger effect in self-evaluation of performance. *Journal of Economic Psychology*, *39*, 85–100. https://doi.org/10.1016/j.joep.2013.07.004

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2013). *Life After P-Hacking*. (S. Botti & A. Labroo, Eds.), *NA - Advances in Consumer Research* (Vol. 41). Duluth, MN: Association for Consumer Research. https://doi.org/10.1086/429607

Spânu, F., Băban, A., Bria, M., Lucăcel, R., Florian, I. Ş., & Rus, L. (2013). Error Communication and Analysis in Hospitals: The Role of Leadership and Interpersonal Climate. *Procedia - Social and Behavioral Sciences*, *84*(0), 949–953. https://doi.org/http://dx.doi.org/10.1016/j.sbspro.2013.06.680

van Zundert, M., Sluijsmans, D., & van Merrienboer, J. (2010). Effective peer assessment processes: Research findings and future directions. *Learning and Instruction*, *20*(4), 270–279. https://doi.org/10.1016/j.learninstruc.2009.08.004

Viswesvaran, C., Schmidt, F. L., & Ones, D. S. (2005). Is There a General Factor in Ratings of Job Performance? A Meta-Analytic Framework for Disentangling Substantive and Error Influences. *Journal of Applied Psychology*, *90*(1), 108–131. https://doi.org/10.1037/0021-9010.90.1.108

Wageman, R., & Gordon, F. M. (2005). As the Twig Is Bent: How Group Values Shape Emergent Task Interdependence in Groups. *Organization Science*. https://doi.org/10.1287/orsc.1050.0146

Willey, K., & Gardner, A. (2010). Investigating the capacity of self and peer assessment activities to engage students and promote learning. *European Journal of Engineering Education*, *35*(4), 429–443. https://doi.org/10.1080/03043797.2010.490577

Williamson, J. M., Lounsbury, J. W., & Han, L. D. (2013). Key personality traits of engineers for innovation and technology development. *Journal of Engineering and Technology Management - JET-M*, *30*(2), 157–168. https://doi.org/10.1016/j.jengtecman.2013.01.003

Wilson, D. S., Ostrom, E., & Cox, M. E. (2013). Generalizing the core design principles for the efficacy of groups. *Journal of Economic Behavior and Organization*, *90*. https://doi.org/10.1016/j.jebo.2012.12.010