Table 1. Surface-level Attributes: Demographic

Attribute	Performance	Moderators	Reference
Gender	Composition (*homogeneous=heterogeneous*) Fisher's Z = -0.38, p > .05	<u>*Overall diversity*</u> Task difficulty *Low* (*homogeneous>heterogeneous*) Fisher's Z= 1.85, p < .05 *Medium* = ns *High* (*homogeneous<heterogeneous*) Fisher's Z= –2.37, p < .01 Task Type Intellectual tasks (*homogeneous=heterogeneous*) Fisher's Z= –2.21, p = 0.99 Performance tasks (*homogeneous>heterogeneous*) Fisher's Z = 3.30, p < .01	Bowers et al., 2000
Bio-demographic diversity	*Performance: Quality* ρ= -.006 [-.09, .08] *Performance: Quantity* ρ= -.02 [-.35, .30] *Social integration* ρ= -.02 [-.08, .04]	<u>*Bio-demographic diversity*</u> Task Complexity – ns Team type – ns Criterion report type – ns Criterion measure type – ns Study setting – ns	Horwitz & Horwitz, 2007
Relations-oriented diversity	*Relations-oriented diversity r = -.03 [-.05, -.02]* *Gender r =-.02 [-.04, .01]* *Occupational demography* *Maj. male setting r = -.09 [-.12, -.05]* *Balanced r = .11 [.06, .15]* *Race/ethnicity r =-.01 [-.04, .01]* *Occupational demography* *Maj. white setting r = -.07 [-.10, -.04]*	<u>*Relations-oriented diversity*</u> Occupational demography (see left column) Industry setting *High-tech r = -.18 [-.20, -.15]* *Service r = -.07 [.05, .09]* *Manufacturing r = -.04 [-.07, -.01]* Interdependence	Joshi & Roh, 2009

	Balanced r = .11 [.07, .14] *Age r =-.06 [-.09, -.04]* *Occupational demography* *Maj. younger setting r = -.08 [-.10, -.05]* *Balanced r = -.05 [-.10, -.00]*	*Low r = .08 [.06, .10]* *Moderate r = -.12 [-.14, -.10]* *High r = -.04 [-.11, .03]* Team type (team duration) *Short-term r = .09 [-.01, .16]* *Long-term r = -.14 [.02, .07]*	
Social categorization diversity	***Performance*** *Overall* ρ= -.024 ***Process*** *Open Com.* ρ= .13 *Freq. of Com.* ρ= -.007	<u>*Social categorization diversity*</u> Uncertainty *Low* ρ= -.046 [-.17, .08] *High* ρ= -.010 Frequency of communication ρ= -.14	Bui et al., 2019
Surface-level similarity	***Process*** *[90%CI]* *Overall information sharing* ρ= .22 [.10, .34] *IS uniqueness* ρ= .27 [.14, .40] *IS openness* ρ= .18 [.02, .34]		Mesmer-Magnus & DeChurch, 2009
Race	*Overall* ρ= -.11 [-.14, -.05] *Lab* ρ= .02 [-.03, .08] *-variety* ρ= .00 [-.06, .07] *Field* ρ= -.13 [-.18, -.06] *-variety* ρ= -.13 [-.18, -.06] *-efficiency* ρ= -.04 [-.07, -.01] *-general perf.* ρ= -.14 [-.20, -.05] *-creativity/innovation* ρ= -.18 [-.34, .01]	Study setting (see left column) Diversity operationalization (see left column) Performance operationalization (see left column	Bell et al., 2011

Sex	*Overall* ρ= -.06 [-.09, -.02] *Lab* ρ= .02 [-.06, .09] -*variety* ρ=.07 [.00, .11] *Field* ρ= -.07 [-.11, -.02] -*separation* ρ= -.01 [-.17, .14] -*variety* ρ= -.09 [-.12, -.04] -*efficiency* ρ= -.09 [-.14, -.03] -*general perf.* ρ= -.06 [-.11, .00] -*creativity/innovation* ρ= -.16 [-.29, -.00]	Study setting (see left column) Diversity operationalization (see left column) Performance operationalization (see left column	Bell et al., 2011
Age	*Overall* ρ= -.03 [-.06, .01] *Lab* ρ= .07 [-.06, .18] *Field* ρ= -.03 [-.07, .01] -*separation* ρ= .04 [-.10, .18] -*variety* ρ= .01 [-.06, .07] -*disparity* ρ= -.04 [-.08, .02]	Study setting (see left column) Diversity operationalization (see left column)	Bell et al., 2011
Demographic diversity	<u>*Creativity/innovation*</u> *Overall demographic diversity* ρ= .01 [-.03, .05] *Gender* ρ= -.04 [-.09, .01] *Age* ρ= .01 [-.04, .06] *Racial/ethnicity* ρ= .03 [-.08, .14] *Educational level* ρ= .00 [-.07, .07]	<u>*Supplemental Analyses*</u> *Demographic diversity* *Country culture* *Collectivistic* ρ= *.05 [-.00, .09]* *Individualistic* ρ= *-.09 [-.17, .00]* *Team collaboration* ρ= -.06 [-14, .03]	Byron et al., 2022
Age	*Composition: Aggregated r= .04 [.01. .07]* *High tech r= .05 [-.07, .18]* *Manufacture r= .06 [-.02, .13]*	*Industry type (see left column)*	Carter et al., 2019

	Service r= .02 [-.03, .08] *Student r= .04 [-.02, .11]* *Composition: Heterogeneous r= -.03 [-.09, .03]* *High tech r= -.22 [-.31, -.13]* *Manufacture r= .06 [-.19, .31]* *Service r= .12 [.05, .19]* *Student r= -.02 [-.11, .06]*		
Race/ethnicity	*Composition: Aggregated r= -.02 [-.10, .06]* *High tech r= na* *Manufacture r=na* *Service r= -.03 [-.14, .08]* *Student r= -.02 [-.15, .10]* *Composition: Heterogeneous r= -.06 [-.11, -.01]* *High tech r= -.19 [-.34, -.04]* *Manufacture r= -.05 [-.13, .04]* *Service r= -.02 [-.13, .08]* *Student r= -.04 [-.11, .03]*	*Industry type (see left column)*	Carter et al., 2019
Gender	*Composition: Heterogeneous r= -.03 [-.05, -.00]* *High tech r= -.07 [-.14, -.00]* *Manufacture r= -.11 [-.18, -.03]* *Service r= -.03 [-.07, .02]* *Student r= .02 [-.03, .07]*	*Industry type (see left column)*	Carter et al., 2019
Faultlines	*Composition: Heterogeneous r= -.05 [-.13, .02]* *High tech r= na* *Manufacture r= -.12 [-.19, -.12]* *Service r= -.05 [-.11, .00]* *Student r= -.11 [-.28, .05]*	*Industry type (see left column)*	Carter et al., 2019
Background diversity	*Innovation* ρ= -.13 [-.32, .05] *Team innovation* ρ= -.10 [-.31, .11]	*Measurement level (see left column)* *Measurement type* *Independent rating* ρ= -.13 [-.33, .08]	Hulsheger et al., 2009
Gender	*Task performance r= -.01 [-.04, .02]* *Gender egalitarianism* *high r= .00 [-.03, .04]*	*Performance type (see below)* *Culture (see left column and below)*	Schneid et al., 2015

	low r= -.07 [-.12, -.01] *Humane orientation* *high r= -.01 [-.04, .03]* *low r= -.03 [-.09, .02]* *Institutionalism collectivism* *high r= -.05 [-.11. .00]* *low r= .00 [-.04, .04]* *Ingroup collectivism* *high r= -.10 [-.18, -.02]* *low r= .00 [-.04, .03]* *Contextual performance r= -.10 [-.18, -.02]* **__Contextual performance__ concerns aspects of an individual's performance, which maintains and enhances an organization's social network and the psychological climate that supports technical tasks (Motowidlo et al., 1997)*	*Objective r= -.02 [-.05, .02]* *Gender egalitarianism* *high r= .00 [-.04, .04]* *low r= -.08 [-.15, -.01]* *Humane orientation* *high r= -.01 [-.06, .03]* *low r= -.04 [-.10, .03]* *Institutionalism collectivism* *high r= -.03 [-.11, .04]* *low r= -.02 [-.07, .02]* *Ingroup collectivism* *high r= -.15 [-.26, -.04]* *low r= -.01 [-.04, .02]* *Subjective r= -.01 [-.06, .03]* *Gender egalitarianism* *high r= .00 [-.05, .06]* *low r= -.05 [-.14, .05]* *Humane orientation* *high r= -.01 [-.06, .04]* *low r= -.02 [-.12, .09]*	

		Institutionalism *collectivism* *high r= -.06 [-.13, .01]* *low r= .02 [-.04, .08]* *Ingroup collectivism* *high r= .00 [-.05, .05]* *low r= -.07 [-.18, .04]*	
Surface-level diversity	*Conflict mES= .08 [.00, .15]* *Comm effectiveness mES= -.16 [-.32, .00]* *Social integration mES= -.06 [-.14, .01]*	*See below section*	Stahl et al., 2009
Surface-level diversity	*Demographic diversity* ρ= -.07 [-.07, -.07]	*80%CI*	Stewart, 2006
Surface-level diversity	*Team performance* *Age r= -.07* *Sex r= -.05* *Racial r= -.03* *Faultlines r= -.14* *Team satisfaction* *Age r= -.09* *Sex r= -.08* *Racial r= -.02* *Faultlines r= -.15*		Thatcher & Patel, 2011*(article retracted)
Demographic diversity	*Overall demographic diversity r= -.02 [-.04, .01]* *Measurement* *objective r= -.01 [-.03, .02]* *subjective r= -.05 [-.08, -.03]* *Rater type* *member r= -.00 [-.05,.05]* *internal leader r= .05 [-.03, .12]* *external leader r= -.06 [-.09, -.03]*	Measurement type (see left column) Task Complexity (see left column) Rater type (see left column) Performance type (see left column)	van Dijk et al., 2012

	Task complexity *low r= -.03 [-.08, .03]* *medium r= -.01 [-.04, .03]* *high r= .01 [-.02, .04]* *Performance type* *in-role r= -.02 [-.04, -.01]* *innovation r= .02 [-.02, .06]*		
Age	*Overall r= -.03 [-.07, .03]* *Measurement* *objective r= -.01 [-.04, .02]* *subjective r= -.10 [-.13, -.07]* *Rater type* *member r= -.01 [-.11, .08]* *internal leader r= .13 [.03, .22]* *external leader r= -.14 [-.18, -.10]* *Task complexity* *low r= -.02 [-.09, .05]* *medium r= -.01[-.07, .04]* *high r= -.00 [-.04, .03]* *Performance type* *in-role r= -.04 [-.05, -.02]* *innovation r= .00 [-.04,.04]*	Measurement type (see left column) Task Complexity (see left column) Rater type (see left column) Performance type (see left column)	van Dijk et al., 2012
Ethnicity	*Overall r= -.05 [-11, .02]* *Measurement* *objective r= -.01 [-.06, .03]* *subjective r= -.14 [-.17, -.11]* *Rater type* *member r= -.08 [-.16, .01]*	Measurement type (see left column) Task Complexity (see left column) Rater type (see left column) Performance type (see left column)	van Dijk et al., 2012

	internal leader r=na *external leader r= -.16 [-.20, -.12]* *Task complexity* *low r= -.03 [-.09, .03]* *medium r= -.01 [-.06, .04]* *high r= .17 [-.03, .35]* *Performance type* *in-role r= -.11 [-.14, -.08]* *innovation r= -.07 [-.16, .02]*		
Gender	*Overall r= -.01 [-.05, .03]* *Measurement* *objective r= -.02 [-.06, .02]* *subjective r= -.06 [-.08, -.03]* *Rater type* *member r= .04 [-.03, .10]* *internal leader r= -.06 [-.16, .05]* *external leader r= -.06 [-.09, -.02]* *Task complexity* *low r= -.04 [-.10, .01]* *medium r= .01 [-.03, .05]* *high r= -.04 [-.09, .02]* *Performance type* *in-role r= -.05 [-.07, -.02]* *innovation r= -.01 [-.07, .06]*	Measurement type (see left column) Task Complexity (see left column) Rater type (see left column) Performance type (see left column)	van Dijk et al., 2012
Nationality	*Overall r= -.01 [-.08, .07]* *Measurement* *objective r= -.08 [-.28, .13]* *subjective r= .00 [-.07, .08]*	Measurement type (see left column) Task Complexity (see left column) Rater type (see left column)	van Dijk et al., 2012

	Rater type *member r= -.03 [-.13, .07]* *internal leader r= -.07 [-.22, .10]* *external leader r= .04 [-.04, .12]* *Task complexity* *low r= na* *medium r= -.01 [-.12, .10]* *high r= .04 [-.07, .15]* *Performance type* *in-role r= .00 [-.07, .08]* *innovation r= -.01 [-.16, .14]*	Performance type (see left column)	
Educational level	*Overall r= .00 [-.06, .07]* *Measurement* *objective r= .02 [-.03, .07]* *subjective r= -.05 [-.08, -.01]* *Rater type* *member r= .02 [-.11, .15]* *internal leader r= .21 [.09, .32]* *external leader r= -.07 [-.12, -.02]* *Task complexity* *low r= na* *medium r= -.14 [-.31, .05]* *high r= .03 [-.02, .08]* *Performance type* *in-role r= -.04 [-.07, -.01]* *innovation r= .20 [.11, .30]*	Measurement type (see left column) Task Complexity (see left column) Rater type (see left column) Performance type (see left column)	van Dijk et al., 2012

Faultlines	*Faultlines r= -.06 [-.17, .05]*		van Dijk et al., 2012
Surface-level diversity	*Creativity/innovation r_c= -.02 [-.11, .06]*	*<u>Surface-level diversity</u>* *Team virtuality* *Collocated r_c=.02 [-.07, .10]* *Non-collocated r_c= -.16 [-.37, .04]* *Task interdependence* *Interdependent r_c= .00 [-.09, .10]* *Independent r_c= -.10 [-.26, .06]* *Task complexity* *Complex r_c= .02 [-.07, .11]* *Simple r_c= -.23 [-.33, -.12]* *Task intellectiveness* *Intellective r_c= -.04 [-.42, .34]* *Judgemental r_c= -.02 [-.10, .06]*	Wang et al., 2019
Less job-related diversity	*Cohesion* ρ= -.03 [.10, -.16] *Performance* ρ= -.07 [.00, -.15]	*<u>Less-job related diversity</u>* *Team type* *TMTs* ρ= -07 [.07, -.22] *Lower-level teams* ρ= -.07 [.02, -.16]	Webber & Donahue, 2001
Relation-oriented variety	*Innovation* *Variety* ρ= -.087 [-.18, .01] *Disparity* ρ= .013 [-.02, .05] *Faultline strength* ρ= -.167 [-.30, -.03]	*Diversity operationalization (see left column)*	Wei et al., 2021
Social category diversity	*Team cognition-performance relationship* *Homogeneous* ρ= .32 [.19, .45] *Heterogeneous* ρ= .42 [.34, .50]	*Note that higher rho means team cogntition is more important to performance*	Niler et al., 2022

Faultlines	*Surface-level social faultlines* *Social interaction quality* ρ= -.072 [-.14, -.01] *Task interaction quality* ρ= -.09 [-.15, -.02]		Zhang & Chen, 2023

Table 2. Surface-level Attributes: Task-related

Attribute	Performance	Moderators	Reference
Ability	Composition (*homogeneous=heterogeneous*) Fisher's Z = -1.13, p > .05	<u>*Overall diversity*</u> Task difficulty *Low* (*homogeneous>heterogeneous*) Fisher's Z= 1.85, p < .05 *Medium = ns* *High* (*homogeneous<heterogeneous*) Fisher's Z= –2.37, p < .01 Task Type Intellectual tasks (*homogeneous=heterogeneous*) Fisher's Z= –2.21, p = 0.99 Performance tasks (*homogeneous>heterogeneous*) Fisher's Z = 3.30, p < .01	Bowers et al., 2000
Task-related diversity	*Performance: Quality* ρ= .13 [.06, .19] *Performance: Quantity* ρ= .07 [.01, .17] *Social integration* ρ= -.04 [-.12, .03]	<u>*Team diversity-Social Integration*</u> Team size Medium ρ= .02 [-.05, .11] Small ρ= -.04 [-.12, .03] Criterion report type Manager/rater-reported ρ= -.19 [-.47, .08] Self-report ρ= -.01 [-.05, .04] Study setting Organization ρ= -.02 [-.09, .02] Laboratory ρ= -.03 [-.09, .06]	Horwitz & Horwitz, 2007
Task-oriented diversity	*Task-oriented diversity r = .04 [.02, .06]* *Function r = .13 [.09, .17]* *Education r = -.02 [-.06, .01]* *Org. Tenure =.03 [-.01, .06]*	<u>*Task-oriented diversity*</u> Industry setting *High-tech r = .06 [.04, .09]* *Service r = -.00 [-.05, .05]*	Joshi & Roh, 2009

		Manufacturing r = .01 [-.05, .06] *Interdependence* *Low r = -.03 [-.11, .06]* *Moderate r = .04 [.02, .06]* *High r = .10 [.05, .15]* *Team type* *Short-term r = .08 [-.01, .16]* *Long-term r = .04 [.02, .07]* *Occupational gender demography* *Maj. male setting r = .06 [.03, .09]* *Balanced r = -.03 [-.08, .02]* *Occupational race/ethnicity demography* *Maj. white setting r = .04 [.02, .06]* *Balanced r = -.02 [-.17, .13]* *Occupational age demography* *Maj. young setting r = .06 [.03, .09]* *Balanced r = .01 [-.03, .04]*	
KSAs diversity	***Performance*** *Overall* ρ= .043 [-.29, .38] ***Process*** *Open Com.* ρ= .135 [-.16, .43] *Freq. of Com.* ρ= .107 [-.31, .53]	*<u>KSA diversity</u>* *Uncertainty* *Low* ρ= .078 [-.13, .28] *High* ρ= .004 [-.41, .42]	Bui et al., 2019

Functional background	*Overall* ρ=.10 [.04, .15] *Variety* ρ=.11 [.05, .15] -*efficiency* ρ=.03 [-.05, .11] -*general performance* ρ=.12 [.03, .18] -*creativity/innovation* ρ=.18 [.02, .30]	Performance operationalization (see left column Team Type -*Design/cross-functional* ρ=.16 *[.08, .20]* -*TMT* ρ=.07 [-.03, .16] -*Other/mixed team type* ρ=-.01 [-.12, .10]	Bell et al., 2011
Educational background	*Variety* ρ= .01 [-.05, .08] -*efficiency* ρ= -.02 [-.04, .01] -*general performance* ρ= -.03 [-.08, .03] -*creativity/innovation* ρ= .23 [.08, .33]	Performance operationalization (see left column -*Team Type* -*Design/cross-functional* ρ= .07 [-.06, .18] -*TMT* ρ= .13 [.03, .21] -*Other/mixed team type* ρ= -.05 [-.13, .04]	Bell et al., 2011
Education level	*Mean* ρ= .01 [-.08, .10] -*intellectual teams* ρ= .11 [-.02, .21] -*physical teams* ρ = -*.07 [one study]* *Diversity* ρ= -.01 [-.05, .04] -*variety* ρ= -.01 [-.09, .08] -*other* ρ= -.01 [-.07, .06]	*Diversity conceptualization (see left column)* *Team type (see left column)*	Bell et al., 2011
Org. tenure	*Mean* ρ= .08 [.01, .13] -*efficiency* ρ= .14 [.09, .17] -*general performance* ρ=.00 [-.11, .11] -*creativity/innovation* ρ= -*.27[one study]* *Diversity* ρ= .04 [-.01, .08] -*separation* ρ= -.03 [-.08, .04] -*variety* ρ= .06 [-.05, .16]	Performance operationalization (see left column *Diversity conceptualization (see left column)*	Bell et al., 2011

	-disparity ρ= .04 [-.01, .10]		
Team tenure	*Mean* ρ= .09 [-.01, .18] *-efficiency* ρ= .11 [.09, .17] *-general performance* ρ= .02 [-.11, .11] *-creativity/innovation* ρ= .10 [-.23, .40] *Diversity* ρ= -.04 [-.10, .02] *-disparity* ρ= -.04 [-.10, .01]	Performance operationalization (see left column *Diversity conceptualization (see left column)*	Bell et al., 2011
Education	*Composition: Aggregated r= -.03 [-.10, .05]* *High tech r= -.10 [-.29, .09]* *Manufacture r= .08 [-.28, .43]* *Service r= .03 [-.08, .13]* *Student r= -.13 [-.26, .03]* *Composition: Heterogeneous r= -.03 [-.08, .01]* *High tech r= -.11 [-.15, -.07]* *Manufacture r= na* *Service r= .06 [-.05, .16]* *Student r= .03 [-.12, .18]*	Industry type (see left column)	Carter et al., 2019
Tenure	*Composition: Aggregated r= .09 [.05, .13]* *High tech r= .03 [-.02, .08]* *Manufacture r= .17 [.10, .23]* *Service r= .05 [.02, .09]* *Student r= .17 [.03, .31]* *Composition: Heterogeneous r= -.00 [-.05, .04]* *High tech r= .08 [.01, .15]* *Manufacture r= .00 [-.07, .08]*	Industry type (see left column)	Carter et al., 2019

	Service r= -.03 [-.10, .03] *Student r= .05 [-.15, .25]*		
Familiarity	*Composition: Aggregated r= .04 [-.02, .11]* *High tech r= .15 [.09, .21]* *Manufacture r= .17 [.08, .26]* *Service r= -.08 [-.16, .00]* *Student r= -.10 [-.18, -.01]*	Industry type (see left column)	Carter et al., 2019
Functional Background	*Composition: Aggregated r= .04 [-.01, .09]* *High tech r= .10 [.04, .17]* *Manufacture r= .42 [.12, .73]* *Service r= -.03 [-.13, .07]* *Student r= .07 [-.01, .15]* *Composition: Heterogeneous r= .06 [.03, .09]* *High tech r= .12 [.10, .15]* *Manufacture r= .07 [-.07, .21]* *Service r= .04 [-.02, .09]* *Student r= -.01 [-.08, .06]*	Industry type (see left column)	Carter et al., 2019
Team Tenure	*Additive* *Performance* ρ= .20 [.15, .24] *Process* *Cognition* ρ= .12 [.00, .23] *Motivational-affective* ρ= .01 [-.03, .05] *Behavioral* ρ= .06 [.03, .10] *Collective*	<u>*Supplemental analyses*</u> *Interdependence* *Additive-Performance* *Low* ρ= .11 [.04, .17] *Med* ρ= .07 [.01, .13] *High* ρ= .29 [.23, .35] *Additive-Cognition*	Gonzalez-Mulé et al., 2020

	Performance ρ= .11 [.05, .16] *Process* *Cognition* ρ= .07 [-.04, .18] *Motivational-affective* ρ= .10 [-.02, .21] *Behavioral* ρ= -.02 [-.06, .02] *Dispersion* *Performance* ρ= .08 [.03, .14] *Process* *Cognition* ρ= .05 [-.01, .11] *Motivational-affective* ρ= .02 [-.05, .09] *Behavioral* ρ= .10 [.03, .17]	*Low* ρ= -.06 [-.23, .11] *Med* ρ= .16 [.07, .25] *High* ρ= .28 [-.15, .70] *Level of Specificity* *Additive-Performance* *Job* ρ= .41 [.35, .47] *Team* ρ= .16 [.12, .22] *Org.* ρ= .10 [.03, .18] *Additive-Cognition* *Job* ρ= .09 [.01, .17] *Team* ρ= .28 [.13, .48] *Org.* ρ= .03 [-.09, .16] *Performance Outcome* *Additive-Performance* *Objective* ρ= .30 [.23, .37] *Subjective* ρ= .07 [.04, .12]	
Job-relevant diversity	*Innovation* ρ= .155 [.00, .31] *Team innovation* ρ= .24 [.04, .44]	Measurement level (see left column) Measurement method *Self-rating* ρ= -.01 [-.16, .14] *Independent rating* ρ= .16 [.02, .30]	Hulsheger et al., 2009
Team longevity	*Innovation* ρ= .02 [-.14, .18] *Team innovation* ρ= -.06 [-.26, .14]	Measurement level (see left column) Measurement method *Self-rating* ρ= -.37 [-.50, -.24] *Independent rating* ρ= .13 [.02, .24]	Hulsheger et al., 2009
Entrepreneurial	*Aggregated Composition r= .14 [.10, .18]*	Industry type (see left column)	Jin et al., 2016

Team Composition - Mixed	*High-tech r= .08 [.03, .14]* *Low-tech r= .21 [.16, .25]* *Heterogeneity r= .05 [.01, .09]* *High-tech r= .07 [.03, .11]* *Low-tech r= -.03 [-.11, .06]*		
Surface-level diversity	*Aggregated* *Expertise* ρ= .16 [.16, .16] Heterogeneity Expertise ρ= -.05 [-.05, -.05] Organizational tenure ρ= -.08 [-.25, .10]	*80%CI*	Stewart, 2006
Surface-level diversity	*Team performance* *Functional r= -.07* *Educational r= -.04* *Tenure r= -.06* *Team satisfaction* *Functional r= -.03* *Educational r= -.05* *Tenure r= -.03*		Thatcher & Patel, 2011*(article retracted)
Job-related diversity	*Job-related diversity r= .05 [.01, .08]* *Measurement* *objective r= .02 [-.01, .06]* *subjective r= .04 [.01, .07]* *Rater type* *member r= .02 [-.06, .09]* *internal leader r= .03 [-.04, .10]* *external leader r= .09 [.06, .13]* *Task complexity* *low r= -.04 [-.09, .02]* *medium r=.03[-.02, .09]*	Measurement type (see left column) Task Complexity (see left column) Rater type (see left column) Performance type (see left column)	van Dijk et al., 2012

	high r=.06 [.02, .09] *Performance type* *in-role r= .04 [.02, .06]* *innovation r= .09 [.04, .14]*		
Functional background	*Functional background r= .07 [.03, .12]* *Measurement* *objective r= .06 [.01, .10]* *subjective r= .12 [.09, .16]* *Rater type* *member r= .13 [.02, .24]* *internal leader r= .15 [.04, .25]* *external leader r= .13 [.09, .17]* *Task complexity* *low r= na* *medium r= .04 [-.04, .12]* *high r= .08 [.04, .12]* *Performance type* *in-role r= .10 [.07, .13]* *innovation r= .13 [.06, .19]*	Measurement type (see left column) Task Complexity (see left column) Rater type (see left column) Performance type (see left column)	van Dijk et al., 2012
Educational background	*Educational background r= -.00 [-.05, .05]* *Measurement* *objective r= -.00 [-.06, .05]* *subjective r= -.00 [-.09, .08]* *Rater type* *member r= .02 [-.16, .19]* *internal leader r= na*	Measurement type (see left column) Task Complexity (see left column) Rater type (see left column) Performance type (see left column)	van Dijk et al., 2012

	external leader r= .06 [-.06, .17] *Task complexity* *low r= na* *medium r= -.01 [-.18, .16]* *high r= .02 [-.05, .09]* *Performance type* *in-role r= -.00 [-.05, .05]* *innovation r= -.00 [-.16, .16]*		
Tenure	*Tenure r= -.01 [-.05, .04]* *Org. tenure r= -.00 [-.07, .06]* *Team tenure r= -.02 [-.07, .04]*	*Measurement type* *objective r= -.01 [-.05, .03]* *subjective r= .01 [-.02, .04]* *Rater type* *member r= -.06 [-.15, .04]* *internal leader r= .00 [-.08, .07]* *external leader r= .08 [.04, .12]* *Task complexity* *low r= -.04 [-.10, .02]* *medium r= -.05 [-.12, .02]* *high r= .01 [-.03, .05]* *Performance type* *in-role r= .01 [-.02, .03]* *innovation r= .03 [-.03, .09]*	van Dijk et al., 2012
Job-related diversity	*Cohesion* ρ= .10 [-.02, .23] *Performance* ρ= .02 [-.03, .07]	*<u>Highly-job related diversity</u>* *Team type* *TMTs* ρ= .03 [-.02, .08] *Lower-level teams* ρ= -.09 [-.22, .04]	Webber & Donahue, 2001
Task-oriented variety *(innovation)*	*Innovation* *Variety* ρ= .11 [.07, .14] *Disparity* ρ= .08 [.04, .12] *Faultline strength* ρ= .11 [.00, .22]	*Diversity operationalization (see left column)*	Wei et al., 2021

Job-related diversity *(creativity and innovation)*	<u>*Creativity/innovation*</u> *Job-related diversity* ρ= .05 [.01, .09] *Functional background* ρ= .06 [.01, .11] *Educational background* ρ= .09 [.01, .16] *Team tenure* ρ= .06 [-.02, .14]	<u>*Supplemental Analyses*</u> *Country culture* *Collectivistic* ρ= *.04 [-.01, .09]* *Individualistic* ρ= *.07 [.00, .13]*	Byron et al., 2022
Functional diversity	*Team cognition-performance relationship* *Homogeneous* ρ= .36 [.17, .56] *Heterogeneous* ρ= .29 [.18, .39]	*Note that higher rho means team cogntition is more important to performance*	Niler et al., 2022
Faultlines	*Surface-level task faultlines* *Social interaction quality* ρ= .08 [-.06, .24] *Task interaction quality* ρ= .155 [.07, .28]		Zhang & Chen, 2023

Table 3. Deep-level Attributes

Attribute	Performance	Moderators	Reference
Personality	Composition (*homogeneous=heterogeneous*) Fisher's Z = -1.2, p > .05	<u>*Overall diversity*</u> Task difficulty *Low* (*homogeneous>heterogeneous*) Fisher's Z= 1.85, p < .05 *Medium = ns* *High* (*homogeneous<heterogeneous*) Fisher's Z= –2.37, p < .01 Task Type Intellectual tasks (*homogeneous=heterogeneous*) Fisher's Z= –2.21, p = 0.99 Performance tasks (*homogeneous>heterogeneous*) Fisher's Z = 3.30, p < .01	Bowers et al., 2000
Personality *Conscientiousness*	*Overall* ρ=.11 [.04, .14] *mean* ρ=.14 [.05, .18] *maximum* ρ=.09 [-.02, .18] *minimum* ρ=.12 [.00, .20] *heterogeneity* ρ=.-.03 [-.12, .06]	Statistical operationalizations (see left column) Study setting Lab ρ= .04 [-.02, .08] Field ρ= .30 [.17, .31]	Bell, 2007
Agreeableness	*Overall* ρ=.12 [.04, .16] *mean* ρ=.17 [.06, .21] *maximum* ρ=.09 [-.04, .18] *minimum* ρ=.19 [.04, .27] *heterogeneity* ρ=.-.04 [-.09, .03]	Statistical operationalizations (see left column) Study setting Lab ρ=.03 [-.03, .08] Field ρ= .31 [.16, .34]	Bell, 2007
Extraversion	*Overall* ρ=.09 [.03, .11] *mean* ρ=.10 [.03, .13] *maximum* ρ=.11 [-.01, .18] *minimum* ρ=.05 [-.04, .12] *heterogeneity* ρ=.03 [-.02, 07]	Statistical operationalizations (see left column) Study setting Lab ρ=.06 [.00, .10] Field ρ= .15 [.06, .19]	Bell, 2007

Emotional Stability	*Overall* ρ=.04 [-.02, .09] *mean* ρ=.13 [.05, .16] *maximum* ρ=.13 [-.03, .25] *minimum* ρ=.07 [-.02, .13] *heterogeneity* ρ=.02 [-.07, .10]	Statistical operationalizations (see left column) Study setting Lab ρ=.03 [-.04, .09] Field ρ= .06 [-.05, .15]	Bell, 2007
Openness to Experience	*Overall* ρ=.05 [-.02,.10] *mean* ρ=.11 [.02, .16] *maximum* ρ=.10 [.01, .15] *minimum* ρ=.05 [-.06, .15] *heterogeneity* ρ= -.03 [-.10, .06]	Statistical operationalizations (see left column) Study setting Lab ρ=.00 [-.06, .07] Field ρ= .20 [.05, .27]	Bell, 2007
Collectivism	*Overall* ρ=.25 [.09, .31] *mean or sum* ρ=.31 [.10, .38] *heterogeneity* ρ=.02 [-.11, .14; lab only]	Statistical operationalizations (see left column) Study setting Lab ρ=.00 [-.09, .08] Field ρ= .35 [.12, .47]	Bell, 2007
Preference for Teamwork	*Overall* ρ=.18 [.02, .29] *mean* ρ=.23 [.08, .32] *heterogeneity* ρ=.01 [-.29, .31; lab only]	Statistical operationalizations (see left column) Study setting Lab ρ=.01 [-.26, .27] Field ρ= .22 [.07, .30]	
Emotional Intelligence	*Mean only* ρ=.18 [.06, .26]	Statistical operationalizations (see left column) Study setting Lab ρ=.20 [.09, .26] Field ρ= .10 [-.22, .39]	Bell, 2007
GMA	*Overall* ρ=.27 [.17, .29] *mean or sum* ρ=.31 [.20, .31] *maximum* ρ=.27 [.07, .37] *minimum* ρ=.34 [.13, .42] *heterogeneity* ρ=.01 [-.07, .09]	Statistical operationalizations (see left column) Study setting Lab ρ=.31 [.19, .29] Field ρ= .18 [.04, .25]	Bell,2007
Personality *Agreeableness*	*Composition: Aggregated r= .14 [.05, .24]* *High tech r= .15 [-.06, .36]*	*Industry type (see left column)*	Carter et al., 2019

	Manufacture r=na *Service r= .40 [.17, .63]* *Student r= .12 [.01, .22]* *Composition: Heterogeneous r= -.14 [-.28, .00]* *High tech r=na* *Manufacture r=na* *Service r= -.24 [-.58, .10]* *Student r= -.12 [-.28, .03]*		
Conscientiousness	*Composition: Aggregated r= .08 [-.00, .16]* *High tech r= na* *Manufacture r=na* *Service r= .15 [.02, .29]* *Student r= .02 [-.09, .12]* *Composition: Heterogeneous r= -.12 [-.25, .01]* *High tech r=na* *Manufacture r=na* *Service r= -.12 [-.29, .05]* *Student r= -.11 [-.50, .28]*	*Industry type (see left column)*	Carter et al., 2019
Extraversion	*Composition: Aggregated r= .10 [.01, .18]* *High tech r=na* *Manufacture r=na* *Service r= .04 [-.09, .17]* *Student r= .13 [.02, .23]* *Composition: Heterogeneous r= .04 [-.10, .18]* *High tech r=na* *Manufacture r=na* *Service r= .04 [-.10, .18]* *Student r=na*	*Industry type (see left column)*	Carter et al., 2019
Openness to Experience	*Composition: Aggregated r= .13 [.02, .25]*	*Industry type (see left column)*	Carter et al., 2019

	High tech r=na *Manufacture r=na* *Service r= .09 [.01, .26]* *Student r= .14 [.02, .24]* *Composition: Heterogeneous r= .14 [-.15, .44]* *High tech r=na* *Manufacture r=na* *Service r=na* *Student r= .14 [-.15, .44]*		
Emotional Stability	*Composition: Aggregated r= .13 [.02, .24]* *High tech r= na* *Manufacture r=na* *Service r= - .00 [-.36, .36]* *Student r= .14 [.02, .27]* *Composition: Heterogeneous r=na* *High tech r=na* *Manufacture r=na* *Service r=na* *Student r=na*	*Industry type (see left column)*	Carter et al., 2019
Cognitive Ability	*Composition: Aggregated r= .07 [-.00, .14]* *High tech r= -.01 [-.19, .17]* *Manufacture r=na* *Service r= .22 [.01, .43]* *Student r= .09 [.02, .16]* *Composition: Heterogeneous r= .07 [-.03, .18]* *High tech r= -.20 [-.42, .02]* *Manufacture r= na* *Service r= .01 [-.19, .21]* *Student r= .10 [.01, .19]*	*Industry type (see left column)*	Carter et al., 2019

Deep-level cultural diversity	*Conflict mES= .05 [-.07, .16]* *Comm effectiveness mES= .14 [-.04, .32]* *Social integration mES= .00 [-.10, .11]*		Stahl et al., 2009
Mixed-cultural diversity *(Surface-level and Deep-level)*	*Creativity mES= .16 [.00, .32]* *Conflict mES= .07 [.01, .13]* *Task conflict mES= .10 [.02, .18]* *Relationship conflict mES= .05 [.03, .13]* *Process conflict mES= .01 [.11, .14]* *Communication effectiveness mES= -.03 [-.15, .09]* *Satisfaction mES= .15 [.05, .25]* *Social integration mES= -.07 [-.12, -.02]* *Performance mES= -.02 [-.04, .00]*	*Task complexity* *-Conflict* *Low mES= -.10 [-.24, .04]* *High mES= .09 [.01, .17]* *Team size* *-Conflict* *Small mES= .12 [.03, .21]* *Large mES= .08 [-.02, .17]* *-Comm effectiveness* *Small mES= .14 [-.07, .36]* *Large mES= -.27 [-.45, -.09]* *-Satisfaction* *Small mES= .28 [.10, .46]* *Large mES= -.04 [-.20, .12]* *-Social integration* *Small mES= -.13 [-.09, .02]* *Large mES= -.17 [-.27, .07]* *Geographic dispersion* *-Conflict* *Collocated mES= .10 [.04, .16]* *Dispersed mES= -.14 [-.32, -.04]* *-Social integration* *Collocated mES= -.08 [-.14, -.02]* *Dispersed mES= .11 [-.03, .25]*	Stahl et al., 2009

		Team tenure *-Conflict* *≤20h mES= .00 [-.10, .10]* *>20h mES= .12 [.05, .20]* *-Comm effectiveness* *≤20h mES= .12[-.06, .29]* *>20h mES= -.14 [-.30, .02]* *-Satisfaction* *≤20h mES= .19 [.07, .30]* *>20h mES=.02 [-.18, .22]* *-Social integration* *≤20h mES= -.12 [-.23, .00]* *>20h mES= -.07 [-.13, -.01]*	
Deep-level diversity	*Overall Deep-level diversity r= -.01 [-.06, .03]* *Personality r= .04 [-.02, .10]* *Extraversion r= .05 [-.05, .16]* *Agreeableness r= -.03 [-.14, .07]* *Conscientiousness r= -.09 [-.19, .01]* *Neuroticism r= .04 [-.09, .17]* *Openness r= .15 [-.00, .30]* *Value r= -.07 [-.18, .04]* *Cognitive r= -.06 [-.16, .03]* *Attitude r= -.04 [-.14, .06]* *Ability r= -.09 [-.29, .13]*	*Deep-level diversity* *Measurement* *objective r= -.01 [-.07, .05]* *subjective r= .00 [-.04, .05]* *Rater type* *member r= -.07 [-.15, .02]* *internal leader r= .01 [-.12. .14]* *external leader r= .01 [-.04, .06]* *Task complexity* *low r= .03 [-.06, .13]* *medium r= .03 [-.02, .08]* *high r= -.18 [-.28, -.08]* *Performance type* *in-role r= -.01 [-.05, .03]* *innovation r= .10 [-.01, .20]*	van Dijk et al., 2012

Deep-level diversity	*Creativity/innovation r_c= .16 [.06, .27]*	*<u>Deep-level diversity</u>* *Team virtuality* *Collocated r_c= .18 [.07, .31]* *Non-collocated r_c= .02 [-.03, .06]* *Task interdependence* *Interdependent r_c= .19 [.10, .30]* *Independent r_c= -.10 [-.43, .23]* *Task complexity* *Complex r_c= .16 [.06, .28]* *Simple r_c= .05 [-.14, .24]* *Task intellectiveness* *Intellective r_c= .09 [-.04, .22]* *Judgemental r_c= .16 [.03, .32]*	Wang et al., 2019
Deep-level diversity	*Task performance r= -.01 [-.03, .01]* *Emergent states r= -.07 [-.10, -.05]* *Task complexity* *Low r= -.00 [-.05, .04]* *High r= -.07 [-.11, -.03]* *Team type* *Exexcutive r= -.22 [-.30, -.14]* *Non-executive r= -.06 [-.09, -.03]* *Team process r= -.10 [-.14, -.07]* *Task complexity* *Low r= -.01 [-.08, .07]* *High r= -.14 [-.19,-.10]*	*Task complexity (see left column)* *Team type (see left column)*	Triana et al., 2021

	Team type *Exexcutive r= -.30 [-.36, -.24]* *Non-executive r= -.02 [-.06, .02]* *Team conflict r= .12 [.07, .16]* *Task complexity* *Low r= .18 [.09, .28]* *High r= .13 [.07, .19]* *Team type* *Exexcutive r= .24 [.13, .35]* *Non-executive r= .09 [.04, .15]*		
Personality diversity	*Emergent states r= -.06 [-.10, -.02]* *Team process r= .03 [-.03, .08]* *Team conflict r= .05 [-.01, .12]*		Triana et al., 2021
Values diversity	*Emergent states r= -.10 [-.14, -.06]* *Team process r= -.26 [-.31, -.21]* *Team conflict r= .16 [.14, .34]*		Triana et al., 2021
Cultural diversity	*Emergent states r= -.02 [-.10, .06]* *Team process r= .03 [-.07, .13]* *Team conflict r= .24 [.14, .34]*		Triana et al., 2021
Deep-level diversity	*Aggreagated* *Personality* ρ= .26 [.12, .41] *Cognitive ability* ρ= .40 [.31, .50]	*80%CI*	Stewart, 2006
Personality	*Elevated [90% CI]* *Extraversion* ρ= .04 [-.05,.13] *Agreeableness* ρ= .24 [.09, .30]	*<u>Team Type</u>* *Elevated* *Agreeableness* *Professional* ρ= .51 [.42,.61]	Peeters et al., 2005

	Conscientiousness ρ= .20 [09, .31] *Emotional Stability* ρ= .04 [-.06, .13] *Openness* ρ= .03 [-.14, .20] *Variability* *Extraversion* ρ= .06 [-.06, .18] *Agreeableness* ρ= -.12 [-.16,-.07] *Conscientiousness* ρ= -.24 [-.33, -.14 *Emotional Stability* ρ= .02 [-.13, .16 *Openness* ρ= -.01 [-.15, .12]	*Student* ρ= .02 [-.11, .15] *Conscientiousness* *Professional* ρ= .42 [.33. .51] *Student* ρ= .00 [-.07, .07] *Emotional Stability* *Professional* ρ= .14 [-.05, .32] *Student* ρ= -.04 [-.07, -.01] *Variability* *Agreeableness* *Professional* ρ= -.13 [-.16, -.11] *Student* ρ= -.08 [-.15, -.01] *Conscientiousness* *Professional* ρ= -.21 [-.34, -.08] *Student* ρ= -.22 [-.36, -.08] *Emotional Stability* *Professional* ρ= .16 [-.01, .33] *Student* ρ= -.11 [-.20, -.02] *Openness* *Professional* ρ= -.11 [-.14, -.08] *Student* ρ= .08 [-11, .26]	
General mental ability	*Operationalization* *Average r = .29 [.23, .36]* *High-member r = .21 [.14, .28]* *Low-member r = .25 [.17, .33]* *Standard Deviation r = -.03 [-.09, .03]*	*<u>General mental ability</u>* *Study-setting* *Lab r = .37 [.32, .42]* *Field r = .14 [.01, .26]*	Devine & Phillips, 2001

Team Orientation	*Team performance* ρ= .46 [.31, .60] *Individual Performance* ρ= -.45 [-.67, -.23] *Conflict* ρ= -.37 [-.49, -.25] *Satisfaction* ρ= .37 [.28, .46] *Innovation/learning* ρ= .84 [.58, 1.09] *Cohesion* ρ= *60 [.44, 76]* *Process* *Communication* ρ= .82 [.65, 1.01] *Coordination* ρ= .78 [.63, 1.01] *Cooperation* ρ= .73 [.63, 1.03] *Trust* ρ= .60 [.41, .79] *Shared mental models* ρ= .69 [.45, .94] *Backup behaviors* ρ= *.72 [.49, .94]*	*Team orientation-team peformance* *Type of team orientation* *Individual-level team orientation* ρ= .23 [.14, .31] *Team-level team orientation* ρ= .53 [.35, .71] *Type of team* *Student* ρ= .20 [.00, .39] *Employee* ρ= .56 [.42, 70]	Kilcullen et al., 2022
Faultlines	*Deep-level social faultlines* *Social interaction quality* ρ= -.13 [-.24, -.03] *Task interaction quality* ρ= -.25 [-.39, -.10] *Deep-level task faultlines* *Social interaction quality* ρ= -.25 [-.35, -.16] *Task interaction quality* ρ= -.28 [-.40, -.16]		Zhang & Chen, 2023