

Willingness to serve is high, but task expectations diverge in disasters

Klaas F. Franzen

`klaas.franzen@uni-luebeck.de`

University of Lübeck, Medical Clinic III

Markus Ries

Heidelberg University

Timo Ziegmann

University of Lübeck, Medical Clinic III

Andreas Batsios

University of Lübeck, Medical Clinic III

Florian Fischer

University of Lübeck, Medical Clinic III

Victoria Edwards

University of Lübeck

Amelie Klapproth

University of Lübeck, Medical Clinic III

Erik Schneider

University of Lübeck, Medical Clinic III

Bernhard Elsner

University of Lübeck

Robert Wunderlich

University, University Hospital Tübingen, Eberhard Karls University

Peter Bradl

German Society for Disaster Medicine (Deutsche Gesellschaft für Katastrophenmedizin e.V.)

Andreas Follmann

German Society for Disaster Medicine (Deutsche Gesellschaft für Katastrophenmedizin e.V.)

Laura Over-Müller

University of Lübeck, Medical Clinic III

Wiebke Zweig

University of Lübeck, Medical Clinic III

Peer Franzen

Reinfeld Volunteer Fire Brigade

Kristina Büscher

University of Lübeck

Daniel Drömann

University of Lübeck, Medical Clinic III

Ole Franzen

Ministry of the Interior, Municipal Affairs, Housing and Sports - Schleswig-Holstein

Kai Mortensen

Cardiology Kiel

Michael Reppel

Cardiology Landsberg

Patrick Ristau

Deggendorf Institute of Technology

Andrea C. Franzen

Schön Klinik Neustadt (i. Holst.)

Maik Forst

Heidelberg University

Research Article

Keywords: surge capacity, willingness to respond, health workforce, disaster preparedness, health professions education, civil protection, cross-sectional survey, Germany

Posted Date: October 19th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10917491/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Personnel surge capacity is a binding constraint of health systems in situations that exceed routine operations. Students and trainees in the health professions are widely regarded as the obvious personnel reserve; however, they are neither systematically registered, nor qualified, nor integrated into alerting and command structures. The extensive literature on willingness to respond almost exclusively surveys the supply-side. Previous research has begun to examine student and organizational perspectives jointly, but task-level concordance between both perspectives remains insufficiently quantified.

Methods

Within the KOMPaKT112-MED/ORG study, we conducted two parallel, anonymous, nationwide online cross-sectional surveys in Germany using item-for-item mirrored instruments: one for students and trainees in health professions (MED) and the other among governing and responsible actors in civil protection and disaster response (ORG). The dyad is constructed at the level of items and constructs, rather than individuals. Therefore, all group comparisons were unpaired comparisons of two independent samples. The primary endpoints were the item-wise endorsement proportions for 11 parallel task areas and their differences between populations. Differences were estimated using Newcombe-Wilson hybrid intervals and adjusted for multiplicity within families using the Holm procedure. Ordinal batteries of unequal lengths were compared using Cliff's delta with bias-corrected and accelerated bootstrap intervals. Rank profiles were compared using Kendall's tau-b with person-level bootstrap resampling. Predictors of willingness were modelled using a cumulative link model with an a priori fixed predictor set, and preference profiles were characterized using latent class analysis. Robustness was assessed across the analysis populations, aggregation rules, exclusion criteria, and field survey waves.

Results

The analysis populations comprised 8,133 MED and 1,656 ORG records (completion rates 87.5% and 69.1%, respectively). Stated willingness among MED respondents and perceived relief potential among ORG respondents were both high: 78.7% of MED respondents were willing to be formally recruited and trained, 84.1% would provide spontaneous, specified and 71.3% of ORG respondents saw substantial relief potential. Nevertheless, ten of the eleven task-area differences were significant after Holm adjustment, with a mean absolute difference of 22.7 pp and a range of -44.0 to 41.7 pp. The rank profiles of the two populations were essentially uncorrelated (Kendall's tau-b = 0.02; 95% CI: -0.02–0.09). The discordance followed a single prespecified content axis: patient-facing acute care was endorsed more often by MED (mean difference + 22.1 percentage points) and care-related, nursing, and logistical tasks more often by ORG (-30.4 percentage points), a difference-in-differences of 52.5 percentage points. Insurance coverage was the only framework condition ranked among the two most important by both populations. Concerns about psychological overload were markedly stronger on the organizational side (Cliff's delta = -0.49; 57.2% versus 22.7%). Prior exposure to

civil protection structures and practical exercises were the strongest predictors of individual willingness. The direction of all 11 differences was preserved across all specifications examined.

Conclusions

Both samples broadly supported the integration of additional personnel, but their task expectations diverged along a clearly identifiable content axis. Concepts for integrating additional personnel should start with the tasks on which both sides already agree, explicitly address expectational gaps in the remaining fields, and include insurance coverage as a central framework condition.

1. Introduction

The personnel surge capacity of a health system is the ability to expand care beyond routine operations at short notice. In the established four-element framework, it is one of four interdependent domains alongside supplies, structures, and systems, and it is the one that cannot be procured, stockpiled, or built ¹⁻³. Where equipment can be purchased and space can be improvised, qualified personnel can only be prepared in advance. This asymmetry becomes decisive in situations that have moved to the center of European preparedness planning since 2022: prolonged power outages, area-wide natural hazards, pandemic and epidemic situations, and states of tension and national defense.

Students and trainees in the health professions are obvious candidates for such reserves. They are numerous, geographically distributed, partially trained in a health occupation, and available at short notice in a way that few other groups are. In Germany, they are neither systematically registered, nor – besides local, limited initiatives – systematically qualified for major incident conditions, nor integrated into alerting and command structures. The national framework directive on total defense adopted in June 2024 foresees the extensive involvement of the civilian health system ⁴, and the federal medical service explicitly names short intensive training courses to establish basic competencies for auxiliary and support activities ⁵. However, an empirical basis for deciding what specific tasks such personnel should be prepared to do is still lacking. The survey of Torlot display a high acceptance for including students into the surge capacity workforce. These results support the need for predefined roles for surge capacity helpers and emphasize the necessity of pre-emptive preparedness training and the clarification of academic obligations ahead of time ⁶.

The international literature offers a large and internally consistent body of evidence on half of this question. Systematic reviews covering three decades and up to 150 primary studies show that willingness to respond among healthcare personnel varies with the type of event, competing obligations towards family, and concerns about personal safety and protective equipment ^{7,8}. A systematic mapping review of the field concluded that research is dominated by single-population designs and that hypothetical willingness is rarely set against any concrete statement of need ⁹. For students, a systematic review of 37 studies covering 11,168 medical students found a weighted mean willingness to volunteer of 68.4%, with a range from 26.7% to 87.8%, alongside poor disaster knowledge and minimal curricular coverage ¹⁰. German and Dutch surveys of students and a nationwide survey of German medical faculties converge on the same picture of an under-implemented curriculum ¹¹⁻¹³.

The perspective of the organizational stakeholders responsible for alerting, briefing, supervising, and deploying such personnel remains largely unexamined. Using comparable instruments, these stakeholders have rarely been asked which tasks they consider suitable for additional personnel. The literature on spontaneous and unaffiliated volunteers demonstrates why this matters. The unprepared convergence of helpers generates substantial coordination, integration, and occupational safety problems, and standards for planning their involvement exist precisely because ad hoc integration reliably fails^{14–17}. If a preparation concept for students and trainees is developed from the supply side alone, it may qualify people for tasks that nobody intends to assign them, while leaving the fields of actual need understaffed.

The present study addresses this gap. We designed two item-for-item mirrored instruments and administered them to the two populations concurrently so that the task willingness of students and trainees could be compared with the tasks that organizational stakeholders considered suitable for additional personnel at the level of individual tasks, competencies, and framework conditions. The aims of the KOMPaKT112-MED/ORG study were 1) to quantify the discordance between the self-reported task willingness and self-assessed competence of students and trainees and the perceived task suitability and competence requirements reported by organizational stakeholders; 2) to test whether this discordance follows a systematic content axis; 3) to identify the framework conditions on which both sides converge; and 4) to determine the modifiable predictors of individual willingness.

We tested the following seven hypotheses, which were specified in advance:

- H1: The item-wise endorsement proportions for the 11 parallel task areas differed between populations.
- H2: The rank order of task preferences differed systematically between the groups.
- H3: Students and trainees endorsed patient-facing acute care tasks more often, whereas organizations endorsed care-related, nursing, and logistic tasks more often.
- H4: Students' willingness was positively associated with prior exposure to civil protection structures and curricular exposure, independent of sociodemographic characteristics.
- H5: Task preferences on the student side are better described by a mixture with at least three latent classes than by a single-class solution.
- H6 (convergence hypothesis): Legal certainty and insurance coverage are rated among the most important framework conditions by both groups.
- H7 (divergence hypothesis) The perceived risk of psychological burden is more pronounced on the organizational side.

2. Methods

2.1 Design

We conducted a non-interventional, anonymous, cross-sectional survey with two parallel, non-individually linked samples. The dyadic character of the design resides at the level of items and constructs, not at the level of persons: the two instruments ask about the same eleven task areas, the same 21 clinical and nursing skills, the same training contents and formats, and the same framework conditions, but from the self-and external

perspectives, respectively. No respondent in one sample corresponded to a respondent in the other. All group comparisons were consequently unpaired comparisons of two independent samples, and procedures for paired data, including limits-of-agreement approaches in their classical form, were neither applicable nor used. This point is stated explicitly because the mirrored construction invites an opposite assumption.

There was no randomization, blinding, intervention, control group in the experimental sense, or interim analysis with consequences for field management. Reporting follows STROBE for cross-sectional studies and CHERRIES for internet-based surveys^{18–20}. Dispositions follow the AAPOR Standard Definitions²¹. Both checklists are provided as supplementary material; the protocol is available online²².

2.2 Populations and eligibility

The MED population comprised persons enrolled in a course of study or vocational training in a health, nursing, therapeutic, or emergency medical service occupation in Germany, covering human and dental medicine, psychology and psychotherapy, nursing, physiotherapy, occupational therapy, speech and language therapy, midwifery, physician assistance, paramedicine, and medical-technical assistants. The ORG population comprised persons holding an organizing, responsible, commanding, or training function in civil protection, disaster response, or public health in Germany, across fire services, emergency medical services, aid organizations, air rescue, the federal agency for technical relief, civil protection authorities, municipal administration, ministries and subordinate authorities, hospitals, the public health service, universities and schools of the health professions, and the federal armed forces medical service.

Both instruments required legal adulthood, sufficient German language proficiency, active consent on the first page, and processing of at least the first substantive question block. Persons under 18 were excluded by design because the subject matter goes beyond opinion polling and addresses willingness to expose oneself to potentially hazardous and distressing deployment situations, including terrorist incidents, chemical, biological, radioactive, and nuclear (CBRN) events, and contact with the dying; involving legal guardians would have required identification of participants and was therefore incompatible with the anonymous design.

2.3 Recruitment and field

Recruitment proceeded exclusively through institutional distribution lists and publicly accessible channels: student representative bodies, university and school distribution lists, professional associations, and learning societies for MED; lower civil protection authorities, leadership of response organizations and aid organizations, state associations, and ministerial distribution lists for ORG. No-named individuals were directly approached by the study team. Onward forwarding by recipients was possible and uncontrolled. Up to two reminder waves were sent through the same channel. The ORG survey opened on July 6, 2026, and the MED survey on July 9, 2026; both surveys closed on August 24, 2026. As the survey was open and no sampling frame existed, no response rate in the strict sense was calculable; we report the completion rate as defined in the corrected version of CHERRIES together with a documented reach estimate^{18,19}.

2.4 Instruments

Both instruments were developed specifically for this project and consented across four development stages by an interprofessional expert panel drawn from emergency and disaster medicine, nursing, emergency medical services, higher education, and civil protection, followed by a cognitive pretest for comprehensibility,

processing time, and filter logic. Both instruments were developed specifically for this study and have not been published elsewhere. The complete instruments, in an English translation of the fielded German version, are provided as Additional File 1 (KOMPaKT112-MED) and Additional File 2 (KOMPaKT112-ORG). Neither the panel feedback nor the pretest responses were part of the analysis population. The instruments were implemented on a LimeSurvey installation hosted in Germany, comprising nine pages and 48 questions each, with median processing times of 13.5 minutes (MED) and 11.8 minutes (ORG); apart from consent and the confirmation of legal adulthood, no item was mandatory, so the completeness check prompted but did not compel a response. They are not externally validated; content validity rests on the expert consenting process and the pretest, and internal consistency and structure are determined here for the first time and reported as exploratory.

Item-for-item parallel batteries exist for the 11 task areas, 21 clinical and nursing skills, 17 assignable training contents, seven training formats, five assignable framework conditions, five assignable barriers or concerns, and the complete availability and mobility block, which uses identical wording in the self-and external perspectives. The full mapping of parallel items across the seven mirrored batteries is given in Supplementary Table S1; the item wording can be verified against Additional Files 1 and 2.

Ordinal batteries occur in four-, five-, and six-point forms. In all six-point batteries, the highest code is a labelled scale endpoint (for example, 'no need at all', 'very unimportant', 'not at all'), not an off-scale category; this was verified against the wording of the consented instruments (see Additional Files 1 and 2 for the full scale labels) and is documented as a protocol clarification. Where MED and ORG batteries differ in the number of steps, the pre-specified hierarchy applies: distribution-free dominance as the primary measure, rule-based collapsing to three levels as the secondary measure, percentile rank as the tertiary measure, and linear normalization as a sensitivity analysis. Differences in linearly normalized means were not reported as primary effect estimates.

2.5 Endpoints

The primary endpoints were the item-wise endorsement proportions for the 11 parallel task areas in both populations and their differences, together with the rank profile of task areas per population and its concordance. Secondary endpoints comprised self-rated confidence in skills among MED respondents compared with the required competence level reported by ORG respondents, formal and spontaneous willingness among MED respondents, the perceived suitability of deploying additional personnel across different situation types as reported by ORG respondents, framework conditions, barriers and concerns in both populations, and availability, mobility, deployment duration, and deployment radius in both populations.

2.6 Statistical analysis

The two-sided significance level was set at 0.05. All estimators are reported with 95% confidence or credible intervals; p-values supplement but do not replace the effect estimation. The full analysis set comprised all cases with consent and at least one substantive item answered with item-wise denominators. The complete set comprised the submitted questionnaires and served as a sensitivity analysis. Three sources of missingness were handled separately: filter-driven non-display, item non-response, and break-off. Dispositions and the break-off distribution for both instruments are reported in Supplementary Table S2.

For the primary analysis, the difference in endorsement proportions was estimated for each of the 11 task areas using the Newcombe-Wilson hybrid method, which combines Wilson score intervals for the two proportions and retains adequate coverage at extreme proportions and unequal denominators ^{23,24}. Multiplicity was controlled using the Holm procedure within the family of task areas ²⁵. The directed contrasts for H3 used the pre-specified cluster assignment (cluster A: basic medical care, assisting physician-led care and pre-triage, patient transport; cluster B: care of uninjured persons, support in long-term care facilities, municipal emergency information points, logistics, registration, and documentation) and were tested one-sided.

Rank concordance was quantified using Kendall's tau-b and Spearman's rho, each with bias-corrected and accelerated bootstrap intervals based on person-level resampling within each sample, so that the uncertainty of the item statistics propagates into the concordance estimate. Because these measures rest on 11 to 21 data points, they are interpreted as estimates with intervals rather than significance tests, and the number of items is reported as the limiting quantity. The full concordance results for all four batteries are given in Supplementary Table S5. A quadrant model dichotomizing item statistics strictly above the respective sample median supplements the concordance analysis, and bootstrap quadrant membership probabilities are reported as a stability measure.

Comparisons across ordinal scales of unequal length used Cliff's delta with bias-corrected and accelerated bootstrap intervals as the primary measure of stochastic dominance ²⁶, supplemented by rule-based collapsing into three levels fixed in advance. Predictors of willingness were modelled using a cumulative link model with a logit link and an a priori fixed predictor set without stepwise selection. The full model output for all three outcome variables is given in Supplementary Table S8. Latent class analysis over the 11 binary task items used a Bernoulli mixture estimated by expectation maximization with 200 random starts per solution and class number selection based on the Bayesian, sample-size-adjusted Bayesian, and consistent Akaike information criteria, together with entropy and interpretability. Exploratory analyses were controlled at a false discovery rate of 5% using the Benjamini-Hochberg procedure and are labelled as hypothesis-generating ²⁷.

Robustness was assessed by varying the analysis population, the aggregation rule for ordinal batteries, the plausibility exclusion rules, and the field wave, and by reporting the sign stability of all 11 primary differences across the resulting specifications. The specification space and the sensitivity analyses are reported in Supplementary Table S11. Analyses were performed in a containerized environment with fixed random seeds; the complete analysis code is available publicly.

2.7 Ethics, data protection and availability

The study was reviewed by the Ethics Committee of the University of Lübeck (reference AZ 2026 – 127, 24th March 2026) in accordance with the Declaration of Helsinki. Participation was voluntary, could be discontinued at any time without disadvantage, and was preceded by active consent as the first question of each instrument. The survey was anonymous in the sense of collecting no personal data: no names, addresses, email addresses, matriculation or personnel numbers were recorded, IP address storage was disabled, no access tokens were used, and regional information was collected only at the level of the first three digits of the postal code. Because no personal reference can be established and re-identification is excluded even with proportionate effort, the data are not personal data within the meaning of the General Data

Protection Regulation (GDPR). Therefore, withdrawal after submission was technically impossible, and participants were informed of this in advance. An incentive draw was conducted through a channel completely decoupled from the survey data.

After publication, the anonymous analysis dataset, the codebook, both survey instruments in their original German version and in English translation, and the complete analysis code will be made available in a public repository (data under CC BY 4.0, code under MIT). Free-text fields are screened for potentially identifying information before release, and cell counts below five are aggregated.

3. Results

3.1 Field, dispositions and samples

The MED instrument recorded 9,375 accessed records, of which 9,033 gave consent, 8,133 were included in the full analysis set, and 7,119 submitted the questionnaire, with a completion rate of 87.5% (Table 2). The corresponding figures for the ORG instrument were 2,055, 1,974, 1,656, and 1,145, with a completion rate of 69.1%. Dispositions for both instruments and the break-off distribution by page are reported in Supplementary Table S2 and Supplementary Figure S1. Both samples exceeded the target sizes stated in the protocol by a wide margin. Break-offs were concentrated on the early pages in both instruments; on the organizational side, a second, smaller peak occurred at the extensive matrix questions. The attrition models identified associations of small effect sizes with occupational groups and organization types; the resulting inverse probability weights were used for sensitivity analysis.

The median processing time was 13.5 minutes for MED and 11.8 minutes for ORG, matching the 12 to 15 minutes stated to participants. The sample characteristics of both populations are presented in Table 1. All postal code lead zones were represented in both samples, with an over-representation of northern Germany.

Table 1

Characteristics of the two analysis populations. The denominators varied by item because filter-driven non-display and item non-response were handled separately.

Characteristic	MED (n = 8,133)	ORG (n = 1,656)
Age, median (IQR), years	24 (22–27)	43 (34–52)
Female, n (%)	4865 (59.8%)	200 (12.1%)
Male, n (%)	3108 (38.2%)	1437 (86.8%)
Largest occupational group / organization type, n (%)	human medicine 2602 (33.0%)	aid organizations 532 (32.4%)
Second largest, n (%)	nursing 2031 (25.7%)	fire services 521 (31.7%)
Third largest, n (%)	emergency medical services 512 (6.5%)	authorities and administration 248 (15.1%)
Late phase of training, n (%)	3930 (48.4%)	not applicable
Active engagement in civil protection or response organizations, n (%)	3226 (40.3%)	not applicable
Prior completed vocational training in health care, n (%)	3874 (47.6%)	not applicable
Employed alongside studies or training, n (%)	4719 (58.0%)	not applicable
At least one practical major incident exercise, n (%)	3503 (45.4%)	not applicable
No real deployment experience yet, n (%)	2369 (30.1%)	not applicable
Primary function level: strategic / administrative, n (%)	not applicable	854 (52.8%)
Primary function level: tactical / operational command, n (%)	not applicable	592 (36.6%)
Command authority in deployment, n (%)	not applicable	1216 (74.7%)
Involved in training or continuing education, n (%)	not applicable	613 (37.7%) regularly; 793 (48.7%) occasionally
Years in a responsible function, median (IQR)	not applicable	14 (5–25)
Major incidents attended in a responsible role, median (IQR)	not applicable	3 (1–8)
Personnel shortages experienced very frequently or occasionally, n (%)	not applicable	875 (82.1%) of 1,066 respondents able to assess

3.2 Both samples broadly support the integration of additional personnel

On the student and trainee side, 80.9% considered the systematic recruitment and training of additional personnel to be sensible or very sensible, 78.7% were willing or definitely willing to be formally recruited and trained, and 84.1% would probably or definitely support the initiative spontaneously if asked in an acute situation. Outright rejection was rare for all three items, at or below 3.5%.

On the organizational side, 82.1% of the 1,066 respondents who felt able to judge reported experiencing personnel shortages in the medical and nursing domains very frequently or occasionally during major incidents, 71.3% assessed the relief potential of additional student personnel as substantial or very substantial, and only 3.0% considered no deployment setting appropriate at all. Willingness to participate in piloting or implementing such a concept was expressed by 41.1% as active or supportive involvement, with a further 44.1% conditional on framework conditions.

These findings indicate broad support for involving additional personnel in both samples. They do not quantify absolute personnel availability or actual organizational staffing needs; the following analyses therefore examine the alignment of task expectations between the two samples.

3.3 Task-level discordance (H1)

Table 2

Endorsement proportions and risk differences for the 11 parallel task areas (MED n = 7,472; ORG n = 1,330). Cluster A denotes patient-facing acute care, cluster B denotes care-related, nursing, and logistic tasks, and items not assigned to either cluster. The intervals are Newcombe-Wilson hybrid intervals; p-values are Holm-adjusted within the family of task areas.

Task area	Cluster	MED %	ORG %	Difference (pp)	95% CI	p (Holm)
Basic medical care / first aid on scene	A	83.2	61.7	21.5	18.8 to 24.2	< 0.0001
Assisting physician-led care / pre-triage	A	73.6	31.9	41.7	39.0 to 44.4	< 0.0001
Patient transport and repositioning	A	66.8	63.6	3.2	0.4 to 6.0	0.0469
Registration and documentation of those affected	B	64.2	80.4	-16.2	-18.5 to -13.7	< 0.0001
Care of uninjured persons and relatives	B	53.3	83.0	-29.8	-32.0 to -27.4	< 0.0001
Psychosocial first support for the affected population	–	31.2	28.5	2.7	-0.0 to 5.2	0.0528
Support in hospitals, admissions, auxiliary wards	–	63.0	74.5	-11.5	-14.0 to -8.9	< 0.0001
Support in long-term care facilities	B	33.9	77.9	-44.0	-46.4 to -41.5	< 0.0001
Municipal emergency information points	B	32.7	70.5	-37.8	-40.4 to -35.1	< 0.0001
Public information and risk communication	–	19.8	37.4	-17.6	-20.4 to -14.9	< 0.0001
Logistic tasks (materials, catering, communication)	B	51.1	75.1	-24.0	-26.5 to -21.4	< 0.0001

H1 is confirmed. Ten of the eleven differences remained significant after Holm adjustment, and the effects were large enough to be practically consequential: the mean absolute difference was 22.7 percentage points, far above the minimum detectable difference of approximately ten percentage points stated in the protocol (Fig. 1, Table 2). The extremes lie in opposite directions and exceed 40 percentage points. Assisting with physician-led care and pre-triage was endorsed by 73.6% of students and trainees but only 31.9% of organizations (+ 41.7 pp), whereas support in long-term care facilities was endorsed by 33.9% of students but 77.9% of organizations (-44.0 pp). Only two task areas showed near agreement: patient transport (+ 3.2 pp) and psychosocial first support (+ 2.7 pp) (Table 2).

The exclusion option was rarely chosen on either side, indicating that the divergence concerned the allocation of tasks rather than the principle of involvement (Supplementary Table S3). Two further task areas were surveyed on the organizational side only, vaccination and infection control assistance, and surveillance and reporting support, and are reported for completeness without entering the discordance analysis (Supplementary Table S3).

3.4 The discordance follows a single content axis (H3)

Table 3

Directed cluster contrasts. The cluster assignment was fixed before the analysis and was not modified in a data-driven manner.

Contrast	Direction specified a priori	Mean difference (pp)	95% CI	p (one-sided)
Cluster A – patient-facing acute care (3 items)	MED > ORG	22.1	20.1 to 24.1	< 0.0001
Cluster B – care-related, nursing and logistic (5 items)	ORG > MED	-30.4	-32.0 to -28.7	< 0.0001
Difference-in-differences A – B (exploratory)	undirected	52.5	49.9 to 55.0	< 0.0001

Both hypotheses were confirmed, and the effects were unusually large (Table 3). Students and trainees endorsed acute care tasks on average 22.1 percentage points more often, and organizations endorsed care-related and logistic tasks 30.4 percentage points more often. The difference-in-differences of 52.5 percentage points quantifies how far apart the two sides stand on the question of what additional personnel are for. This figure is the central quantity of this study.

The same axis reappears in the deployment settings surveyed on the organizational side. Care and shelter facilities (86.8%), general hospital wards (82.5%), and long-term care facilities (75.6%) were considered conceivable by three-quarters or more of the respondents, whereas the pre-hospital incident scene (21.8%) and intensive care or operating theatre settings (25.0%) were named by approximately one-fifth to one-quarter. Organizations conceive of additional personnel as a rearward and inpatient resource, not as a resource at the scene.

3.5 Rank concordance differs sharply between domains (H2)

Table 4

Rank concordance of item profiles between the two populations. Bootstrap intervals were based on person-level resampling within each sample. Given the small number of items, these measures were interpreted as estimates with intervals and not as significance tests.

Battery	k items	Kendall's tau-b (95% CI)	Spearman's rho (95% CI)
Task areas	11	0.02 (-0.02 to 0.09)	0.06 (0.03 to 0.10)
Skills (confidence versus requirement)	21	0.70 (0.67 to 0.73)	0.85 (0.83 to 0.89)
Training contents	17	-0.29 (-0.34 to -0.26)	-0.42 (-0.45 to -0.36)
Training formats	7	0.62 (0.52 to 0.71)	0.82 (0.75 to 0.89)

Concordance was strikingly heterogeneous across the four parallel batteries (Fig. 2, Table 4). For task areas, there was effectively no agreement (tau-b = 0.02, interval including zero; rho = 0.06). For skills, agreement is high (tau-b = 0.70): both sides prioritize the same competences and differ in the competence levels required by ORG respondents or self-reported by MED respondents rather than in the ranking of the competences. For training content, the rank correlation was negative (tau-b = -0.29), meaning that the two sides pulled in

opposite directions. Training formats occupied an intermediate position ($\tau\text{-}b = 0.62$) (Fig. 2, Table 4, full results in Supplementary Table S5).

This pattern is informative. The two populations converge when reasoning about concrete clinical competencies and diverge when reasoning about roles, responsibilities, and curricular priorities. Any interpretation that treats discordance as uniform would obscure this.

3.6 Self-reported confidence and required competence levels differ across several skills

Self-rated confidence on the student side (six-point scale) was set against the competence level required by organizations (four-point scale) for 21 parallel skills. Following the pre-specified hierarchy, the primary comparison was the proportion of respondents at the upper end of each scale (all 21 skills in Supplementary Table S4)., that is, students reporting high confidence and organizations requiring at least safe independent application (Table 5).

Table 5

The four largest gaps in each direction across 21 parallel skills. A positive gap means that the requirement exceeds reported confidence, and a negative gap means that reported confidence exceeds the requirement. The complete table is provided in Supplementary Table S4.

Skill	MED high confidence, %	ORG requiring safe application or more, %	Gap (pp)
Self-protection at the scene	61.6	90.0	28.4
Hemorrhage control: pressure dressing	69.9	88.5	18.6
Rescue breathing (mouth-to-mouth or mouth-to-nose)	56.9	73.2	16.3
CPR with an automated external defibrillator	70.1	85.6	15.5
Structured assessment (cABCDE or MARCH)	64.0	42.7	-21.3
Empathetic communication	79.9	70.5	-9.4
Airway management	55.5	48.0	-7.6
Clinical observation	67.4	61.5	-6.0

The largest single gap is not in an invasive procedure but in self-protection at the scene (28.4 pp), a competence that is a precondition of any deployment, and that organizations also rank first among the desired training content. Hemorrhage control, rescue breathing, and defibrillation follow. In the opposite direction, reported confidence exceeds the stated requirement most clearly for structured assessment schemes (-21.3 pp): these are taught in health professions curricula but are not expected of additional personnel by the organizations that would deploy them (Fig. 3, Table 5).

3.7 One convergence and one divergence (H6, H7)

H6 was only partially confirmed. Insurance cover satisfied the convergence condition in every bootstrap replicate; it lies among the two most important framework conditions in both populations (bootstrap probability 1.00) and is rated highly important by 94.7% of organizations. Legal certainty is important on both sides but does not reach the top two positions on the student side, where it is displaced by supervision from experienced instructors, rated as highly important by 95.6% of respondents. The full results for all five mapped framework conditions are given in Supplementary Table S7. The joint condition stated in H6, both attributes in the top two of both groups, is therefore not met (bootstrap probability: 0.00). For concept development, this means that insurance coverage is the undisputed minimum condition, whereas the second most important condition is population specific.

H7 was clearly confirmed and constituted the largest single divergence among the barriers. Of the organizations, 57.2% expressed strong concern about the psychological overload of additional personnel, compared with 22.7% of students and trainees naming this as a strong personal barrier (Cliff's delta = -0.49; 95% CI -0.51 to -0.46). This is notable because the direction is reversed for other barriers: for insurance and availability, it is the students and trainees who voice the stronger reservations. Therefore, concept development must provide psychosocial preparation and aftercare, even though the target group does not register this need with comparable urgency.

Another organizational result deserves mention because it contradicts a widespread assumption. Concern about competition with or delimitation from established forces was the weakest of all nine concerns, rated strongly by only 14.7% of respondents. The leading concerns were psychological overload (57.2%) and additional command and supervision burden (57.1%) (Supplementary Table S7). The obstacle is a resource problem, not a cultural one (Fig. 4).

3.8 Exposure, not socialization, predicts willingness (H4)

In the cumulative link model for formal willingness ($n = 7,235$), the strongest predictor after adjustment was active engagement in civil protection or response organizations (OR 2.49; 95% CI 2.12–2.94), followed by practical exercise experience (OR 1.78; 1.54–2.05), former engagement (OR 1.45), and radio communication training for law enforcement and emergency services (OR 1.32). Curricular exposure contributed independently but weakly, with an odds ratio of 1.06 per standard deviation and a confidence interval within the pre-specified region of practical equivalence. Occupational group differences were substantial, with nursing and emergency medical service trainees showing lower adjusted willingness than the human medicine reference group, and midwifery and physician assistant students showing higher willingness. The results for spontaneous willingness were materially identical (Supplementary Table S8).

Therefore, H4 is confirmed in its central claim, with an important restriction: the modifiable component of the association, curricular exposure, is weaker. We report this plainly because it supports the case for curricular reform less strongly than initially expected.

3.9 Preference profiles are heterogeneous (H5)

Information criteria decreased monotonically up to the six-class solution, and entropy remained above the pre-specified target of 0.70 throughout; therefore, H5 is confirmed: a single-class solution describes task

preferences considerably worse than a mixture with at least three latent classes. The six-class solution was reported for interpretation, with class prevalence ranging from 11.3% to 23.9%. Model selection and class prevalences are given in Supplementary Table S9, and the item-conditional response probabilities in Supplementary Figure S2.. The classes differed less in the overall level of willingness than in their profiles: alongside a broadly deployable type and a type focused on acute care, a distinct type emerged that would accept only care-related and administrative tasks. An analogous analysis of organizational deployment-profile classes is reported in the supplement and juxtaposed with the student classes substantively, not statistically, since psychometric equating between the two instruments is not admissible in the absence of shared persons or identically scaled items.

3.10 Robustness

The primary quantity was stable across all specifications examined (Supplementary Table S11). Varying the analysis population between the full analysis set, the completer set, and the full analysis set excluding flagged straight-liners, varying the aggregation rule for ordinal batteries, varying the plausibility exclusion rules, and excluding the incentive-weighted first field week left both the direction and the order of magnitude of the discordance unchanged, and all 11 task-area differences retained the sign observed in the primary analysis in every variant. Therefore, the substantive conclusion does not depend on individual analytical decisions. The weighting of the official marginal distributions foreseen in the analysis plan could not be executed because the required external reference distributions were unavailable; this is reported as a protocol deviation and discussed under limitations.

4. Discussion

4.1 Principal findings

To our knowledge, this study provides the first national comparison of task willingness among health profession students and trainees with the tasks that organizational stakeholders consider suitable for additional personnel in major incidents and disaster response. Both samples expressed broad support for the involvement of additional personnel: more than three-quarters of the student and trainee respondents were willing to be formally recruited and trained, and more than two-thirds of the organizational respondents perceived substantial relief potential. However, the rank profiles of the eleven parallel task areas were effectively uncorrelated, and the largest differences exceeded 40 percentage points in both directions. These findings do not quantify absolute personnel availability or actual organizational staffing needs; instead, they identify a substantial discordance in task expectations between the two samples.

The fit problem is not diffused. It follows a single content axis that was specified in advance and confirmed with an unusually large effect. Students and trainees orient themselves towards patient-facing acute care, organizations towards care-related, nursing, and logistic support. The difference-in-differences of over 50 percentage points is, in our view, the most consequential number this study produces because it translates directly into the risk that a preparation concept developed from either side alone would systematically miss the other.

4.2 Why the two sides look in different directions

Both perspectives are internally consistent. Students and trainees' reason from their prospective professional role: What they are learning and expect to do is the clinical care of individual patients. Organizations reason from their operational reality: In prolonged, area-wide situations, the binding constraints are shelter facilities, general wards, and long-term care, not the incident scene, where established forces are trained, equipped, and command-integrated. Neither of these views is incorrect. They collide because they have never been related to one another.

The same axis is visible across the four independent parts of the instruments, which is the strongest internal argument that it is not an artifact. It appears in the task areas, deployment settings, and competence requirements, where structured assessment schemes are the skills on which reported confidence, most exceeds the stated requirement, and in the training content, where the rank correlation is negative. A one-off divergence in a single battery could be explained by item wording; however, a consistent reversal across four batteries constructed independently of one another cannot.

The most serious alternative explanation is a construct asymmetry inherent in the mirrored design: The student instrument asks about willingness to participate oneself, the organizational instrument about realistic deployment of others, and these are not identical constructs. We take this objection seriously and have reported it as a limitation. However, it does not account for the observed pattern. A pure construct difference would generate a uniform level shift, with one side systematically endorsing more than the other, across all items. We observe a crossing pattern, with each side endorsing more in one cluster and less in the other. Level shifts and crossings are distinguishable, and only the latter is compatible with substantive differences in task conception.

4.3 What the two sides do agree on

Where the two populations converge, they converge on protection rather than content. Insurance coverage is the only framework condition that lies among the two most important for both groups in every bootstrap replicate. Legal certainty is important on both sides but is displaced on the student side by supervision from experienced instructors, which is the single highest-rated framework condition in that population. This is a practically useful asymmetry. The organizational side asks for a legal and command basis, the student side for guidance and reassurance, and both ask for insurance.

The high-ranking concordance for clinical skills points in the same direction. Both sides consider the same competencies important; however, they differed in the competence levels required by organizational respondents and the confidence levels self-reported by students and trainees. Concordance in ranking and discordance in level are distinct phenomena and conflating them misrepresents the findings. Notably, the competence gap is largest for skills that rank high in both orderings, with self-protection at the scene at the top. These findings support giving self-protection a prominent position in a preparation curriculum.

4.4 Implications for concept development

Three implications can be stated with reasonable confidence.

1. **Starting from the consensus tasks, but not stopping there.** Registration, documentation, and support in hospitals and auxiliary wards are endorsed above the respective median by both populations and can be

translated into a concept immediately. Tasks with an expectation surplus, above all support in long-term care facilities, care of uninjured persons, and logistics, require active recruitment work because the need exists, but the willingness does not. Tasks with a willingness surplus, above all assisting physician-led care and pre-triage, require explicit expectation management, since high willingness meets a markedly reserved organizational assessment. A concept that quietly assigns students to the consensus quadrant without addressing the surplus in either direction will lose the group it recruits: the acquisition of practical competence is the strongest motivator on the student side, and a purely administrative role does not deliver it.

2. **Make self-protection and psychosocial preparation non-negotiable curriculum elements.** These are the two areas in which the organizational side sees the greatest deficit and the greatest risk, and in which the student side does not register a comparable need. Self-protection showed the largest competence gap of all 21 skills, and psychological overload showed the largest divergence of all barriers. In both cases, the target group's self-assessment cannot serve as a needs assessment.
3. **Insurance coverage was treated as the entry condition and supervision as the retention condition.** Supervision by experienced personnel is the highest-rated condition on the student side and the eighth of ten on the organizational side, which is precisely the kind of asymmetry that produces implementation failure if left unexamined: the resource the target group considers indispensable is one the deploying organizations rank comparatively low, while simultaneously naming the supervision burden as their second strongest concern. This tension cannot be resolved by exhortation and requires an explicit staffing decision.

4.5 Relation to existing evidence

The willingness measured in our study estimates are consistent with those in the international literature. A systematic review by Byrne et al. reported a weighted mean willingness to volunteer of 68.4% across 37 studies and 11,168 medical students, with a range of 26.7% to 87.8%¹⁰. Our figure of 78.7% lies above the pooled mean but well within the reported range, and the self-selected nature of an open-field survey is a plausible explanation for the position within that range. The same review documents a significant discrepancy between planned and actual volunteering, which is why we describe all our estimates as stated willingness^{28,29}.

Our curricular findings, reported in full in a companion analysis, reproduce the picture established for German medical students by Wunderlich et al. nine years after the approval of the national concept curriculum and confirmed from the faculty perspective by Kassermann et al.^{11,12}. Comparable findings from the Netherlands and Italy suggest that this is not a peculiarity of Germany^{13,30}.

Our findings depart from expectations in the organizational reception of additional personnel. The German literature on emergent citizen responses describes disaster myths as impediments to collaboration between unaffiliated responders and professionals¹⁵. Our data show that competition and delimitation are the weakest of the nine concerns. We do not read this as a contradiction but as a difference in the object. The literature concerns spontaneously converging, unknown, and unprepared helpers. Our instrument asked about prepared, professionally pre-trained personnel integrated in advance. The reservation appears to be related to

unpreparedness rather than to the person, which is an encouraging finding for concept development and a testable proposition for future work.

Finally, the profile of situations in which ORG respondents considered the deployment of additional personnel appropriate, reported in detail in a companion analysis, placed states of tension and national defense at the top, with a substantial margin over prolonged power outages and pandemic situations, while everyday mass casualty incidents ranked lowest. This pattern contrasts with a literature that remains largely focused on mass casualty incidents^{1,7} and aligns with the direction of current German civil defense policy^{4,5}. Because the survey was conducted in mid-2026, this pattern may be time-bound, and we report the field period for that reason.

4.6 Strengths and limitations

The principal strength of this study is its design. Two item-for-item mirrored instruments administered concurrently to both sides of the same question permit a fit analysis at the item level that single-population designs cannot produce. The interprofessional breadth across seven occupational groups and the inclusion of situation types up to and including the state of national defense extend the evidence base into areas that are, as far as we can establish, empirically unoccupied. The sample sizes considerably exceeded the planned figures, the analysis code was frozen before application to the real data, and the primary quantity was stable across all examined specifications.

However, several limitations should be considered when interpreting these results.

- All results concern stated willingness and not observed behavior. The intention-behavior gap is well documented, and a systematic review of student volunteering empirically shows the discrepancy for this population [9,26,27]. Therefore, our willingness figures should be read as an upper bound.
- The cross-sectional design did not permit causal statements. While curricular exposure is associated with willingness, this does not establish that additional teaching produces willingness; self-selection into relevant offerings is equally plausible, and the two cannot be separated here.
- Both samples were self-selected and drawn without a sampling frame; therefore, no response rate in the strict sense is calculable. The proportion of respondents with prior civil protection affiliation on the student side is a selectivity indicator, and its effect is examined in the pre-specified subgroup analyses. The weighting of the official marginal distributions foreseen in the analysis plan could not be executed because the reference distributions were unavailable. Duplicate participation was prevented by browser-based cookies, but because IP addresses were not stored and no access tokens were used, repeat entries from a different browser or device cannot be excluded; this is an accepted trade-off of the anonymous design.
- The construct asymmetry between the self and external perspectives discussed in Section 4.2 is inherent in the design and cannot be eliminated, only argued around.
- The instruments were not externally validated. Psychometric properties were determined for the first time and are reported as exploratory. Psychometric equating between the two skill batteries is inadmissible in the absence of shared persons or identically scaled items; therefore, the juxtaposition of confidence and requirement is substantive rather than metric.

- The concordance measures rested on 11 to 21 data points and were interpreted as estimates with intervals rather than significance tests.
- This was a single-country study conducted within one health and civil protection system. We make no claim of transferability, although the framing within the European and Western alliance context suggests that the questions generalize, even if the answers may not.
- The task questions were not anchored to a single standardized disaster scenario. Respondents may therefore have considered different event types, durations, and operational conditions when evaluating the tasks, which may partly account for the observed differences between the two samples.

A parallel national survey of medical students on training, confidence, and participation in crisis and disaster situations is being conducted independently in Germany. It addresses a single occupational group from the supply side; the present study addresses seven occupational groups and both sides of the fit question. We regard two bodies of evidence as complementary when they can be related.

4.7 Directions for further research

Three questions were posed. First, whether the discordance observed here narrows when both sides are brought into structured contact, which is testable in a before-and-after design around joint exercises. Second, whether the association between prior exposure and willingness survives a longitudinal design, which would separate the selection interpretation from the formation interpretation, is unclear. Third, and most demanding, whether stated willingness translates into actual participation, which requires either a registry of prepared personnel with subsequent activation data or an exercise-based behavioral endpoint. Until such data exist, the intention-behavior gap remains the binding limitation of this literature, including ours.

5. Conclusions

Both stated willingness among health profession students and trainees and perceived relief potential among organizational respondents were substantial. However, the two samples showed marked task-level discordance. This discordance was not diffuse, it followed a content axis between patient-facing acute care and care-related, nursing, and logistic support, was robust to the analytical choices examined, and was visible across four parts of the instruments.

From a systems thinking perspective, preparedness is ultimately determined not by how individual actors or organizations perform in isolation, but by how effectively they work together across organizational boundaries to ensure coordination, trust, fairness, and continuity of care³¹. Concepts for integrating additional personnel should therefore begin with the tasks on which both sides already agree, work actively in the fields in which need exceeds willingness, manage expectations explicitly where willingness exceeds need, and treat insurance coverage as the one non-negotiable precondition. The findings also support including self-protection and psychosocial preparedness in any training concept. The reservation that concept developers most commonly anticipate, resistance from established responders, is not what the data show; the binding constraint on the organizational side is the supervision and command burden, which is a resource question and therefore potentially solvable.

Declarations

Ethics approval and consent to participate

The study was reviewed by the Ethics Committee of the University of Lübeck (reference AZ 2026 – 127, 24th March 2026) in accordance with the Declaration of Helsinki. Participation was voluntary and preceded by active informed consent as the first question of each instrument. The survey did not collect direct identifiers such as names, email addresses, personnel numbers, or IP addresses.

Consent for publication

Not applicable.

Competing interests

The authors declare no conflicts of interest

Funding

This study was funded by the institutional resources of Medical Clinic III, University of Lübeck. There was no external funding, industry involvement, or commercial interest.

Author Contribution

KFF, MRi, and TZ contributed equally and share first authorship. PR, ACF, and MvdF contributed equally and share last authorship. Conceptualization: KFF, MRi, TZ, VE, RW, PB, PF, ACF, MvdF. Methodology: KFF, MRi, TZ, BE, RW, PB, AFo, PR, MvdF. Software: TZ, PR. Formal analysis: KFF, TZ, PR. Validation: MRi, BE, RW, AFo, KM, MRe, PR. Investigation: KFF, AB, FF, VE, AK, ES, LOM, WZ, PF, KB, OF, ACF. Resources: FF, PF, KB, DD, OF. Data curation: KFF, TZ, AB, AK, ES. Writing – original draft: KFF, TZ, MRi, PR, MvdF. Writing – review and editing: all authors. Visualization: TZ, KFF. Supervision: KFF, MRi, DD, PR, ACF, MvdF. Project administration: KFF, LOM, WZ. Funding acquisition: KFF, DD. All authors read and approved the final manuscript.

Acknowledgement

We thank the expert panel of the instrument consenting process, the pretest participants, and the institutions that forwarded the survey invitation, subject to their consent to be named.

Data Availability

The data will be provided upon request.

References

1. Barbisch DF, Koenig KL. Understanding Surge Capacity: Essential Elements. *Acad Emerg Med*. 2006;13:1098–102.
2. Kaji A, Koenig KL, Bey T. Surge Capacity for Healthcare Systems: A Conceptual Framework. *Acad Emerg Med*. 2006;13:1157–9.
3. Hick JL, Koenig KL, Barbisch D, Bey TA. Surge Capacity Concepts for Health Care Facilities: The CO-S-TR Model for Initial Incident Assessment. *Disaster Med Pub Health Prep*. 2008;2:S51–7.
4. Bundesministerium des Innern. *Rahmenrichtlinien Für Die Gesamtverteidigung - Gesamtverteidigungsrichtlinien - (RRGV)*.
https://www.bmi.bund.de/SharedDocs/downloads/DE/veroeffentlichungen/themen/sicherheit/RRGV.pdf?__blob=publicationFile&v=4 (2024).
5. Bundeswehr. *Gesundheitsversorgung in Der Landesverteidigung*.
<https://www.bundeswehr.de/resource/blob/6001406/badda37108b51bf7c8ee191c391ce0b2/bericht-symposium-gesundheitsversorgung-in-der-landesverteidigung-data.pdf> (2025).
6. Torlot L, et al. [Medical students in hospital emergency preparedness in the wake of the COVID-19 pandemic: a qualitative analysis of the ESCAPE study]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*. 2026;69:666–75.
7. Chaffee M. Willingness of Health Care Personnel to Work in a Disaster: An Integrative Review of the Literature. *Disaster Med Pub Health Prep*. 2009;3:42–56.
8. Santinha G, Forte T, Gomes A. Willingness to Work during Public Health Emergencies: A Systematic Literature Review. *Healthcare*. 2022;10:1500.
9. Nafar H, et al. A systematic mapping review of factors associated with willingness to work under emergency condition. *Hum Resour Health*. 2021;19:76.
10. Byrne MHV, Ashcroft J, Alexander L, Wan JCM, Harvey A. Systematic review of medical student willingness to volunteer and preparedness for pandemics and disasters. *Emerg Med J*. 2022;39:e6.
11. Wunderlich R, et al. Self-Perception of Medical Students' Knowledge and Interest in Disaster Medicine: Nine Years After the Approval of the Curriculum in German Universities. *Prehosp Disaster Med*. 2017;32:374–81.
12. Kasselmann N, Willy C, Domres BD, Wunderlich R, Back DA. Implementation of disaster medicine education in German medical schools - a nationwide survey. *GMS J Med Educ*. 2021;38:Doc79.
13. Mortelmans LJM, et al. Dutch senior medical students and disaster medicine: a national survey. *Int J Emerg Med*. 2015;8:34.
14. Twigg J, Mosel I. Emergent groups and spontaneous volunteers in urban disaster response. *Environ Urbanization*. 2017;29:443–58.
15. Lorenz DF, Schulze K, Voss M. Emerging citizen responses to disasters in Germany. Disaster myths as an impediment for a collaboration of unaffiliated responders and professional rescue forces. *J Contingencies Crisis Manag*. 2018;26:358–67.
16. Nielsen LR. Proposing a model for integrating spontaneous volunteers in emergency response in Denmark. *Int J Disaster Risk Reduct*. 2024;108:104533.

17. International Organization for Standardization. Security and resilience – community resilience – guidelines for planning the involvement of spontaneous volunteers.
18. Eysenbach G. Improving the Quality of Web Surveys: The Checklist for Reporting Results of Internet E-Surveys (CHERRIES). *J Med Internet Res*. 2004;6:e34.
19. Eysenbach G, Correction. Improving the Quality of Web Surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). *J Med Internet Res*. 2012;14:e8.
20. von Elm E, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet*. 2007;370:1453–7.
21. The American Association for Public Opinion Research. Final Dispositions of Case Codes and Outcome Rates for Surveys. (2023).
22. Franzen KF et al. Einsatzbereitschaft Studierender und Auszubildender der Gesundheitsberufe und organisationaler Bedarf im Bevölkerungsschutz: eine dyadisch-parallele Querschnittserhebung in Deutschland (KOMPaKT112-MED/ORG) – Analyse- und Publikationsplan. *protocols.io* <https://doi.org/dx.doi.org/10.17504/protocols.io.j8nlk7do1g5r/v1> (2026)
[dx.doi.org/10.17504/protocols.io.j8nlk7do1g5r/v1](https://doi.org/dx.doi.org/10.17504/protocols.io.j8nlk7do1g5r/v1)
23. Newcombe RG. Interval estimation for the difference between independent proportions: comparison of eleven methods. *Stat Med*. 1998;17:873–90.
24. Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat Med*. 1998;17:857–72.
25. Holm SA. Simple Sequentially Rejective Multiple Test Procedure. *Scand J Stat*. 1979;6:65–70.
26. Cliff N. Dominance statistics: Ordinal analyses to answer ordinal questions. (1993).
27. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J Roy Stat Soc: Ser B (Methodol)*. 1995;57:289–300.
28. Sheeran P, Intention–Behavior Relations. A Conceptual and Empirical Review. *Eur Rev Social Psychol*. 2002;12:1–36.
29. Sheeran P, Webb TL. The Intention–Behavior Gap. *Soc Pers Psychol Compass*. 2016;10:503–18.
30. Ingrassia PL, Ragazzoni L, Tengattini M, Carenzo L, Della Corte F. Nationwide Program of Education for Undergraduates in the Field of Disaster Medicine: Development of a Core Curriculum Centered on Blended Learning and Simulation Tools. *Prehosp Disaster Med*. 2014;29:508–15.
31. Ries M, Ribeiro Pombo MH, Behan D, Welker A. Who Leads When No One Is in Control? A Systems Thinking Approach to Navigating Complexity in European Population Health, Conflict and Hybrid Crises. SSRN Scholarly Paper at <https://doi.org/10.2139/ssrn.7205498> (2026).

Figures

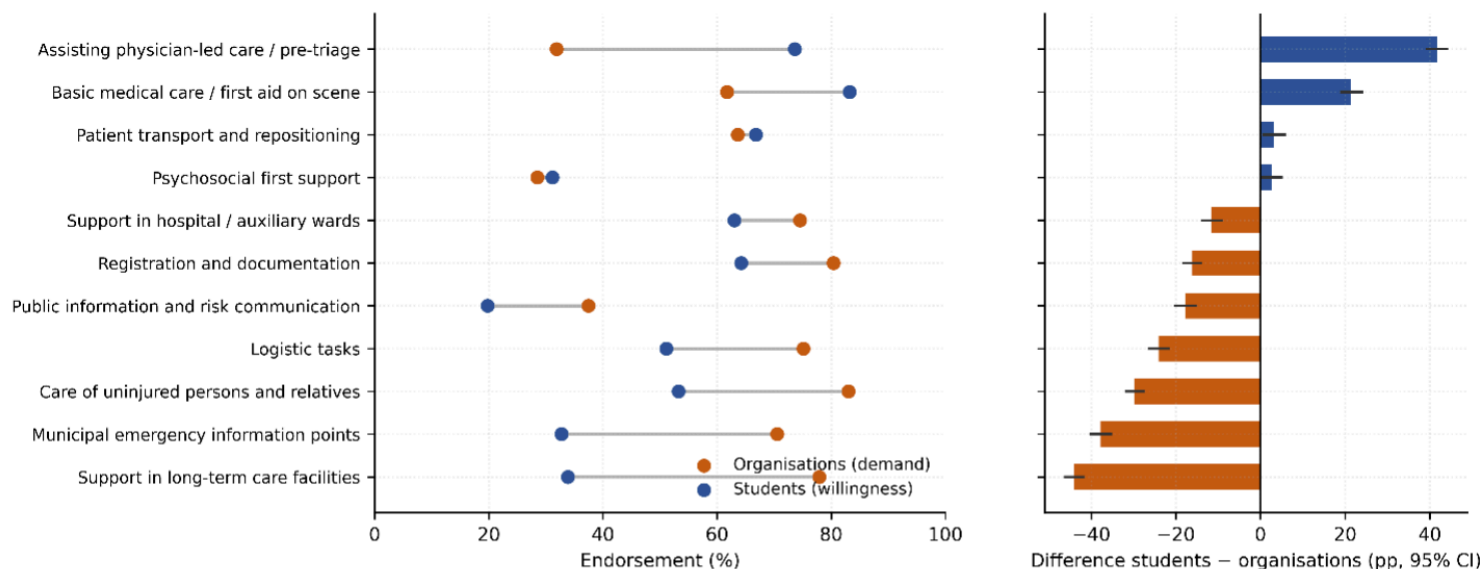


Figure 1

Task-level discordance across eleven parallel task areas. The left panel shows the endorsement proportions of both populations connected by a line, and the right panel shows the risk difference with its 95% confidence interval. Positive values indicate a willingness surplus; negative values indicate an expectation surplus.

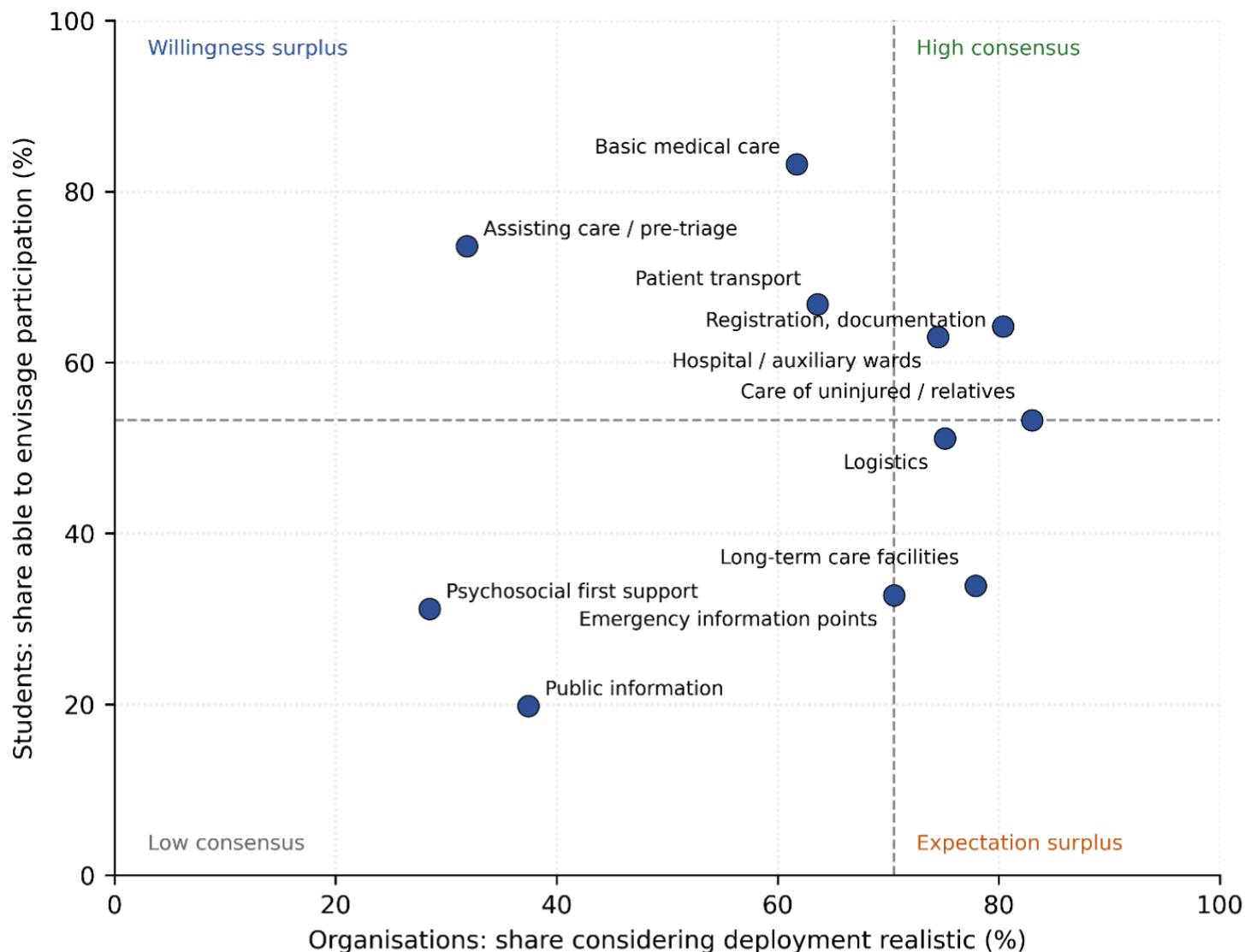


Figure 2

Quadrant model of the eleven task areas, with axes dichotomized strictly above the respective sample median. Two items lie exactly at the median of their distribution, and their assignment is therefore rule-dependent; this is documented in Supplementary Table S6.

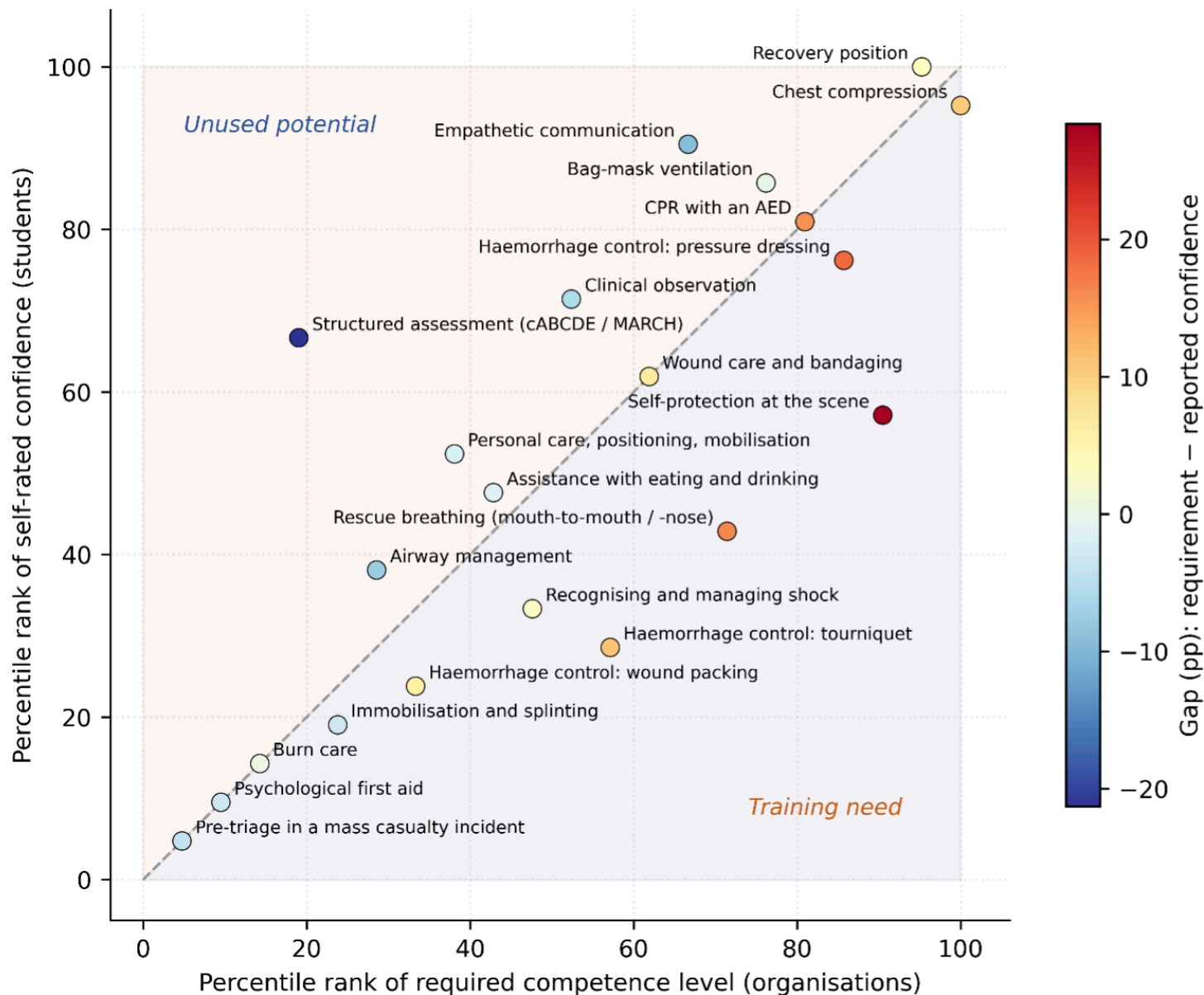


Figure 3

Competence gap matrix across 21 parallel skills, contrasting self-rated confidence among students and trainees with the competence levels required by organizations.

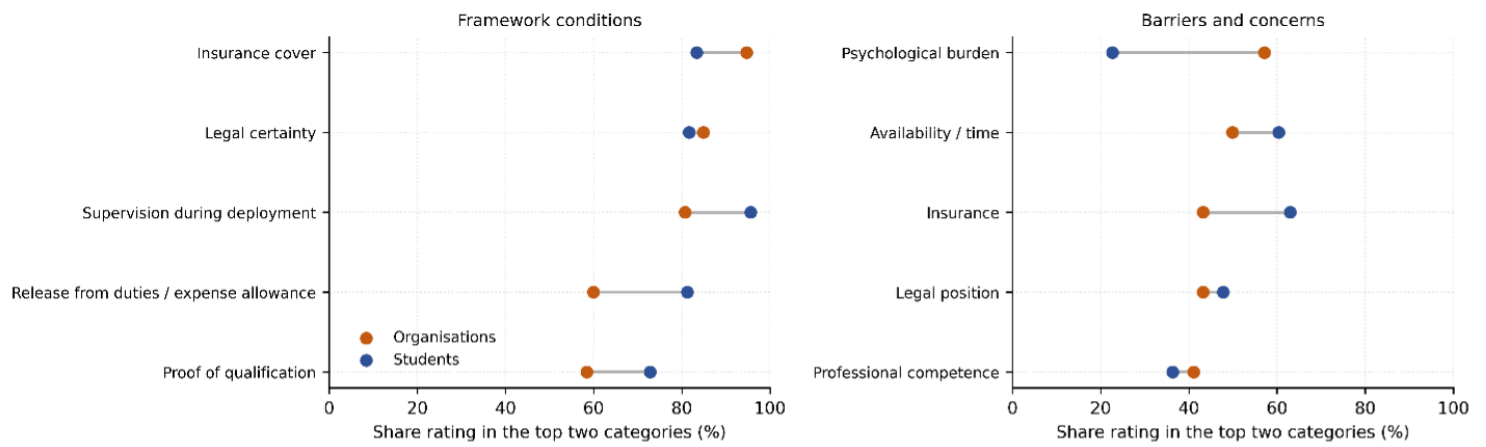


Figure 4

Framework conditions (left) and barriers or concerns (right) in both populations, shown as the share rating each item in the top two categories.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [20260903Publication1Kompakt112studS2V1.docx](#)
- [20260903Publication1Kompakt112studS1V1.docx](#)
- [20260908KOMPakt112P1SupplementaryMaterialcombinedV1.docx](#)