Funded by the European Union

INEGMA-E²

International Network of Evaluators & Guideline for a Methodological Approach in Exercise Evaluation

D4.1 Skill-Set for Evaluators

WP4 – Network of Evaluators

Dissemination Level: Public

Deliverable Type: R(eport)

Date: Jan, 31th, 2023

Author/s: Sandra Pfister (DCNA); René Kastner (DCNA); Bernhard Bürger (AIT); Dr. Celia Norf (BBK);

DCNAustria Disaster Competence Network Austria

Federal Office of Civil Protection and Disaster Assistance

LAUREA AMMATTIKORKEAKOULU University of Applied Sciences

AIT AUSTRIAN INSTITUTE OF TECHNOLOGY TOMORROW TODAY

iTTi Universität der Bundeswehr München

DOCUMENT INFORMATION PAGE	
Grant Agreement No. 101047665	Deliverable Due Date: 31 January 2023
Project Starting Date: 01 January 2022 (18 months)	Actual Submission: 06 February 2023
Deliverable Number: D4.1 (WP4)	Leading Partner: DCNA

KEY WORDS
skill set, competence requirements, evaluators, evaluation experts

AUTHORS & CONTRIBUTORS		
Author	**Institution**	**Entire Document-Section-Appendix**
Sandra Pfister	DCNA	Entire Document
Bernhard Bürger	AIT	Methodology, Results
Celia Norf	BBK	Entire Document
Contributors		
René Kastner	DCNA	Methodology
Christian Resch	DCNA	Methodology
Vlatko Jovanvoski	DCNA	Methodology
Gerald Schimak	AIT	Methodology
Lisanne Steininger	BBK	Methodology
Jari Savolainen	LAUREA	Methodology
Hendrik Bruns	Uni BW	Methodology
Samuel Zientek	ITTI	Methodology

REVIEWS		
Reviewer	**Institution**	**Entire Document-Section-Appendix**
Hannes Kern	DCNA	Entire Document
Hendrik Bruns	Uni BW	Entire Document
Gerald Schimak	AIT	Entire Document

VERSION HISTORY		
Release	**Status**	**Date**
0.1	Initial Draft	25 Jan 2023
0.2	Internal Review	26 Jan 2023
0.3	Consortium Review	26 Jan 2023
0.4	Second Draft	31 Jan 2023
0.5	Internal Final Review	04 Feb 2023
1.0	Final Version	06 Feb 2023

Disclaimer

Disclaimer

Table of Contents

List of Figures

Executive Summary

The main goal of this report is to elaborate on a skill-set for evaluators, which support the recruitment process of exercise evaluators and contributes to quality assurance for future exercise evaluations in context of civil protection.

The deliverable gives a short introduction into the general approach to skills and competencies (1), and the applied methodology (2), before it focuses on the identified skills and competence requirements that are required for exercise evaluators (3). Finally, it ends with some concluding remarks (4).

About this project

In the context of civil protection exercises, well-considered and extensive evaluation plays a crucial role in documenting the best practices and shortcomings recognized during the exercises' conduction. By identifying lessons learnt evaluation is essential for a constant improvement in training efforts, thus promoting the capacities of response units in the European Union and its neighbouring countries for dealing with real disaster scenarios. The project INEGMA-E² is building upon an upcoming approach of independent evaluation and aims for a new level of exercise evaluation in context of civil protection, which will meet high standards concerning documentation, replicability, and goal orientation.

The three pillars of the project are: 1) The development of an adequate and versatile evaluation methodology, addressing the different types of existing exercises. Each of those has different needs and goals, thus requiring diverse evaluation approaches. 2) Exploring the great number of existing tools, which can facilitate the data collection throughout the exercise process. Software solutions and

technical tools like databases and hand helds empower the evaluators to collect a great amount of data even under difficult circumstances as part of the training reality. 3) The creation of an international pool of evaluators, which will be accessible by all institutions managing those kinds of exercises, to ensure the availability of highly skilled experts when needed. Those invited to this pool of evaluators will have to meet a certain skill set developed during the project.

A strong interconnection of all three essential fields – methods, tools and network – is crucial for setting new standards in civil protection exercise evaluation. By ensuring the provision of results for future exercises INEGMA-E² will significantly contribute to a continuous improvement of exercise outcomes. In addition, it will connect experts with exercise evaluation, will create a mechanism to share knowledge and good practices and will be designed for further grow and scale up of evaluation quality.

About this deliverable

This deliverable identifies skills and competence requirements for exercise evaluators, and by doing so it serves as basic document for the conceptualization of a training for evaluators (T4.2) and the concept for the network of evaluators (T4.3). For one thing, it derives training priorities, for another thing, it derives cornerstones for the competence profiles for the subscribed evaluators and the recruitment process.

The report was elaborated in cooperation of DCNA, AIT and BBK, and with valuable contributions from the project consortium.

Abbreviations and Glossary

A common glossary of terms for all INEGMA-E² deliverables, as well as a list of abbreviations, will be made available at the INEGMA-E² website.

1 Towards a competence profile for exercise evaluators

Exercise evaluation has long been a relatively neglected topic in scientific literature (Beerens et al. 2020). Generally, the term evaluation includes both the output of an evaluation (i.e., the evaluation report), and the highly demanding process of assessing evaluation objects (Beerens 2020), following various concepts, e.g., in terms of goals such as development, accountability or knowledge and/or the focal point such as systems, structures or processes (Bruns et al. 2022), and methodologies (Ruoslahti and Lonka 2022). In the following report, disaster evaluation refers to the processual dimension and includes the whole process from planning to documenting and presenting evaluation findings.

Exercise evaluation is of high importance for the long-term benefit of civil protection exercises, which are often high-demanding in terms of resource input. It is considered to fulfill the critical functions of sharing experience and avoiding mistakes, contributing to learning and improvement (Beerens et al. 2020b). Against this background, there is a need for a shared and solid understanding of options and how to best conduct evaluations.

While some researchers worked on the development and elaboration of tools and methods to support the evaluation process, little has been said about the "best way to conduct and present the output" (Beerens et al. 2020). Topics such as the quality assurance of the evaluation, compliance with certain standards, codes of conduct or the adequacy of applied methods and methodologies are rarely examined. Previous findings from projects Meta-evaluation, i.e. the "evaluation of evaluation" – and, one may add, the evaluators – Beerens and Thaler (2016) conclude, seems to be a relative blind spot in the literature on disaster exercise evaluation. Against this background, there was growing interest in the research on how to plan and conduct evaluation. For example, Beerens et al. (2020), elaborated a guideline for the evaluation description, or examined the practitioners' expectations from disaster evaluation.

Regarding "the evaluators", there still seems to be a lack of literature. Little has been said about the requirements for evaluators. Some practical guidelines and handbooks for exercise organizers and/or evaluators mention some requirements, such as the need for external evaluators, or the need to train evaluators on the specific exercise, e.g., in terms of the participating actors etc. (e.g., Swedish Civil Contingencies Agency 2011, 2016). The INEGMA-E² project contributes to that gap. It analyses the set of skills and knowledge needed for a solid evaluation of disaster exercises and develops a competence profile for disaster evaluators.

The competence profile for disaster evaluators is informed by the competence model for the Union Civil Protection Mechanism Training Programme suggested by Karapidis and Dworschak (2016, 24 et segg.). Using a workshop at DG ECHO, they identified four generic competence classes used to define the profiles needed in crisis management organisation. However, the model was not applied to the competence profiles of exercise evaluators.

In the following, *knowledge* is understood as a body of information, facts, theories, etc., while *skills* or *know-how* describe the ability to do something. *Competence* refers to a more abstract level and is a combination of various skills and knowledge that enable individuals to achieve certain goals within a given context (here: exercise evaluation). It is, so to say, about bringing knowledge and skills into action within a given situation. This is also the reason, why terminology in the following report alternates ("skill-set" and "competence profile"). While the plan was to define a skill-set of evaluators, it became very clear that skills (knowledge and know-how) alone not make up a good evaluator, but rather, they must be able to make use of their knowledge and know-how within the highly demanding context of civil protection exercises.

We use the term *competencies* to refer to single knowledge areas or skills/knowhow that, together, make up a certain competence.

The first competence class suggested by Karapidis and Dworschak (2016) was *professional competence.* This includes skills and knowledge of organizations, processes, tasks and workplaces. Usually, it consists of a body of facts, theories and practices and is an outcome of learning.

Second, *methodological competence* includes skills that can be applied to different contexts (ibid.). Within the INEGMA-E² project it is understood as both the know-how to apply tools and methods needed to evaluate exercises and more general practical skills and competencies that enable individuals to master complex situations and navigate through the exercise context. Thus, it is suggested to distinguish between *methodological competencies* (in a narrow sense), and *practical skills* and more *"basic" requirements.*

Third*, social competence* enables individuals to act in socially interactive situations. It includes a range of (predominantly, "soft-")skills that enable individuals to communicate or collaborate with others (ibid.).

Fourth, *personal competence* is the ability for individual development (ibid.). While the other competence classes relate to content or other actors, personal competence somehow relates to the attitude towards oneself and personal traits.

2 Methodological Approach

As it was not possible to identify a model of requirements for exercise evaluators, the INEGMA-E² study has an exploratory character. Its purpose is the development of a set of skills and knowledge, and, as a result, a competence profile for exercise evaluators. To do so, it conducts a delphi expert panel, which elaborates on an informed consent among experts, rather than a representative aggregation of individual opinions.

Data was gathered through computer-based, semi-structured questionnaires (see Annex). Two panels have been conducted. In the first wave, panellists were invited to an initial assessment of skills and competencies required for evaluators. In the second wave, respondents were invited to comment on the preliminary results.

The sampling strategy was a conscious sampling, consisting of researchers (17 identified via Scopus Literature Review) and practitioners (evaluators in former ModEx or FSX, self-recruitment during Nicosia Risk Forum 2022) in the field of exercise evaluation. As the purpose of a delphi survey is an informed consent among the experts, rather than statistical representation, thus a small sample size about 10-20 (5-10 scientific experts + 5-10 practical experts) was pursued. 61 have been invited to participate in the first wave (66 were identified, 5 invalid contact details). As an incentive, participants were offered the opportunity to pre-register for the network of evaluators. A total of n=20 responded to the questionnaire, with n=19 having completed the questionnaire, which is a response rate of 31%.

Regarding the roles of the respondents, 37% are researchers in the field of disaster evaluation, 68% are exercise evaluators, 74% are trainers and/or organizers, 74% have participated in an exercise as trainee and 10% have other roles related to exercise evaluation (e.g., project manager). The sample is rich in the diversity of countries of origin, each located in Europe despite one from Israel. There is a strong gender imbalance in the sample with just 21% of female participants. Respondents are on an average age of 49, with the youngest being 33 and the oldest being 67. 58% declare an affiliation to a practical organisation, e.g., including emergency operation centres, civil protection authorities on various scales, humanitarian organizations or first responder organizations.

Participants were asked to agree on the temporary storage of their data during the duration of the study and local storage with accessibility being restricted to the project partners. This consent is important for the longitudinal study, as wave 2 refers to answers given in wave 1, thus requiring a personalized access key. Further, participants were asked at the end of the questionnaire, if they want to pre-register for the network of evaluators[1], and agree to store their email-address for this purpose in an external file that cannot be linked to the response data, once the study is closed.

The questionnaire in the first wave was a semi-structured questionnaire, i.e., a combination of closed and open questions. Participants were invited to rate the priority of seven sets of items, derived from previous deliverables, protocols from discussions and further input from the project consortium and literature. For one thing, they were directly derived, e.g., when “language skills” were mentioned as requirement for evaluators during the workshops conducted in WP3. For another thing, they were indirectly derived, e.g., when it was mentioned during the workshops, that, usually, *teams* of evaluators were deployed for the exercise evaluation, “cooperativeness” was derived as a potential skill. Further, they were provided to opportunity to make comments, e.g., on the circumstances or additional skills.

In the second wave, the participants (n=20) were asked to comment on the preliminary results from wave 1 (on tendencies, as well as consent or dissent) and were invited to change their initial assessment made in wave 1. Thus, the questionnaire consisted of statements derived from the

[1] The experts willing to participate will be invited to the online platform directly after the launch of the “Network of Evaluators.”

preliminary findings and the opportunity to comment on these statements. The response rate was at 40% (n=8) with varying degrees of item non-response. A decrease was expected for qualitative and in-depth comments.

Data analysis was oriented towards finding an informed consent among experts. Thus, an analysis of descriptive indicators together with a qualitative analysis of open comments is appropriate. Assuming that the distances on the response scales are approximately equal, descriptive parameters for the items were examined. The mean and mode values were taken as an indicator for the average assessment within the expert sample, the standard deviation was taken as an indicator for consent or dissent.

As the sampling strategy for the delphi panel was not oriented towards statistical generalizability, but rather was a conscious sample that aims for a set of rich expertise, it is important to explicitly emphasize that the values provided below are just valid for the groups of panellists and do not represent the opinion of the civil protection community. This is in accordance with the overall purpose of the delphi panel that is to find an informed consent on the required skills and competencies required for exercise evaluators.

To do so, the survey conducters examined descriptive parameters (mode, mean, standard deviation - assuming constant “distances” between not required, low, moderate, and high priority) and put strong emphasize on the open comments of both rounds. The plots provided below provide information on the central tendency (Y-Axes), taken as an indicator for the perceived priority of certain skills and competencies, and scattering (X-Axes), taken as an indicator for the consent or dissent of the participating experts. The graphs are to be interpreted as follows: Items (skills/competencies) that are located on the left-bottom corner are rated, on average, with high priority and high consent. These were interpreted as *universal requirements*, which each evaluator shall feature. Items that are located in the upper part are rated with relative low priority. These features were interpreted as nice add-ons, that may, indeed, be relevant in specific situations, but not to be taken as pre-requirements for exercise evaluators. Items that are located in the middle were rated with moderate priority and neither consent nor dissent. These items resemble skills and competencies that either have high priority but just for certain tasks and contexts, or as skills and competencies that evaluators should have, but not as necessary preconditions for becoming an evaluator. This is in line with the remark of a respondent, namely that high priority may indicate a requirement for all evaluators, while lower priority may add value to the evaluation teams, but varies along with the roles/tasks and type of evaluation. Another one argued, lower priority does not necessarily mean that the competence at hand would not add value to the evaluation at all, but that some of the competencies may contribute more, and some complement each other. Finally, items that are located on the right-side in the middle, resemble skills and competencies with extreme dissent, so to say a bifurcation of the expert group with one group assigning high priority and the other low priority to not required (see figure 1). It is important to note that discriminatory power is relatively low. Rather, tendencies and distances between items are to be considered.

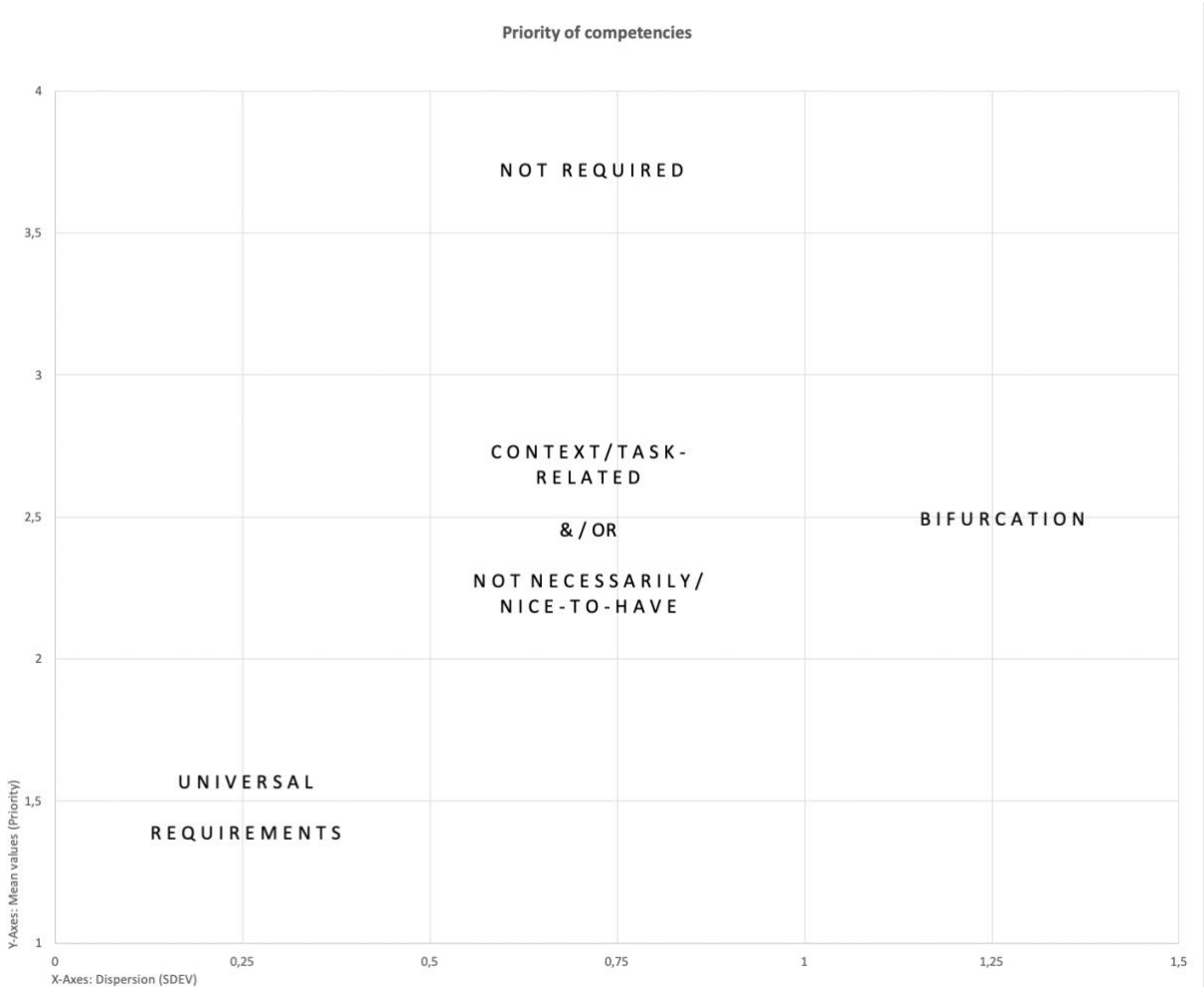

Figure 1: Priority of skills and competencies

3 Results

The following section provides the results of the delphi study.

Figure 2 gives an overview over all 106 items. A broad set of skills and competencies is, on average, rated with moderate priority and medium deviation. Beyond that, a smaller set of skills and competencies has been rated with high priority on a rather consensual basis (i.e., lower dispersion). Few have been rated with no to low priority and some have been rated with very high dissent.

In this regard, it is interesting that items for professional competence and knowledge (orange) have been rated with lower priority than items for other competence areas. None of them may be considered a universal (or “standard” requirement) other than some key, personal and social competencies. Rather, knowledge on specific subject matters may highly depend on the planned exercise and its context.

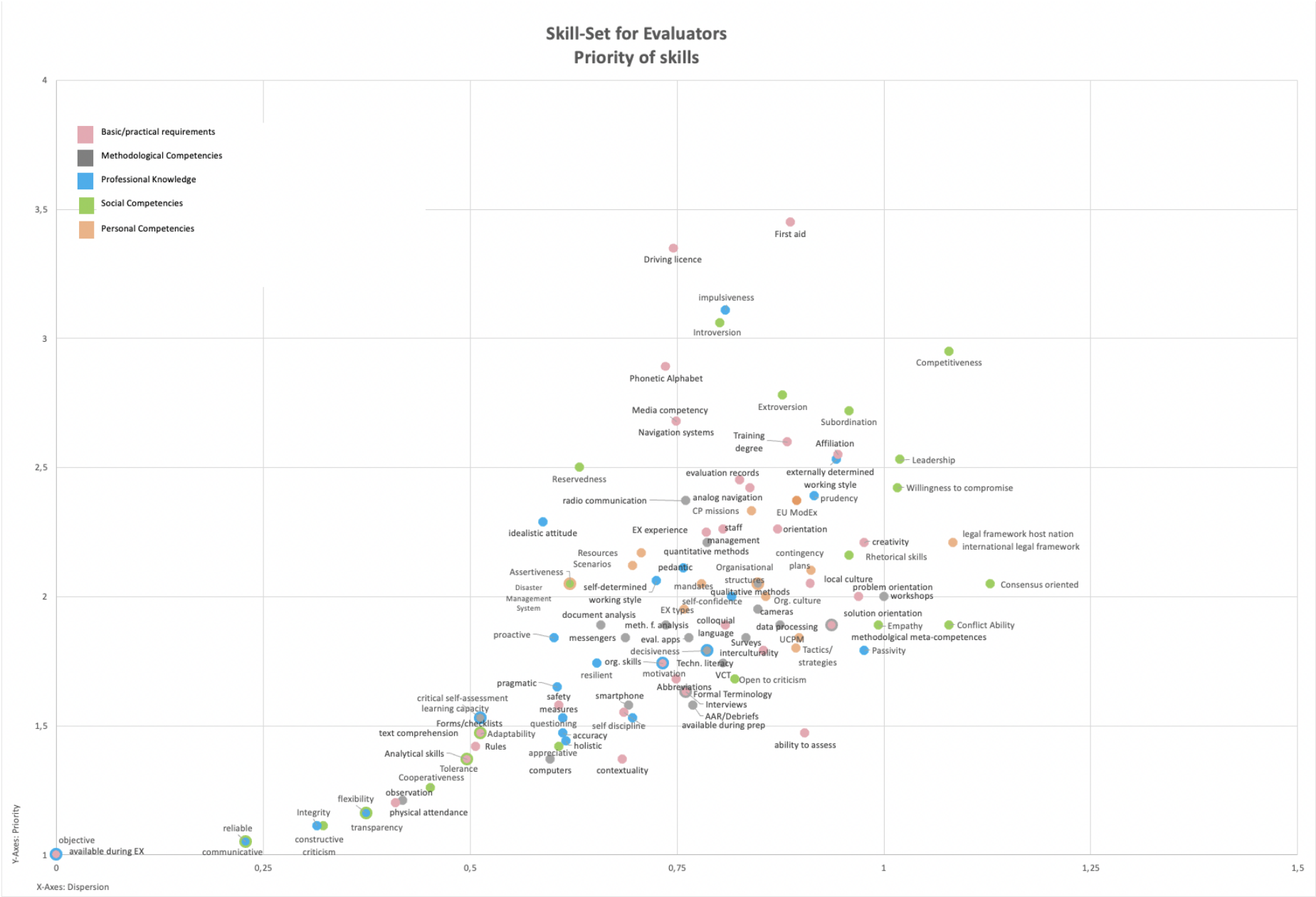

Figure 2: Prioritization of skill

3.1 Basic Requirements and Practical Skills

- **High priority/consent**: Availability[2] during exercise, and physical attendance were rated with high priority and high consent. Availability during exercise planning/preparation was also rated with relatively high priority, but still lower than availability and physical attendance during the exercise. Further, analytical skills, contextuality and text comprehension were rated with high priority and relative consent. Also, issues that relate to proper behaviour during the exercise were rated with high priority and consent such as the knowledge of the safety measures, and the rules of engagement.
- **Moderate-high priority/neither nor-dissent**: Competencies that enable understanding and comprehension were also prioritized, e.g., the familiarity with colloquial language, and the knowledge of formal terminology and abbreviations, and intercultural competence (moderate to high priority). Further, ability to assess and organisational skills were rated with moderate to high priority. Ability to assess was rated with relatively high priority and with higher dispersion.
- **Moderate priority/low consent**: the knowledge of local culture, as well as creativity, problem and solution orientation were rated with moderate priority and lower consent.
- **Moderate to low priority**: Indicators for practical experience, such as previous exercise participation, evaluation records, an evaluation training degree as well as the affiliation to first responder organisations were rated with moderate to low priority and lower consent. The same applies to competencies that enable to navigate through the scenery, such as spatial awareness/orientation and the capability to apply analog navigation or navigation systems. Staff management capabilities were also rated rather low. The same applies to the ability to properly communicate with the media and the mastery of the phonetic alphabet.
- **Low priority-not required**: Finally, the possession of a driving licence and the ability to provide first aid were rated with very low priority or were not required at all.

We have aggregated and interpreted the results as follows:

Involvement: The involvement in the exercise is a basic requirement of outstanding importance. In the second round of the study, each respondent agreed upon that (n=7). However, exceptions may exist in specific cases. Evaluators must be at least available during the exercise. Preferably, they are also available during the planning phase, but this is not a prerequisite. One respondent argued, "*Evaluators indeed could be good to participate in the planning phase, but it is not necessary and it is often not the case as they are appointed at a later stage." (Rd. 2).* Usually, it should be sufficient, if the lead evaluator is available then, *"if the evaluators work as team, availability during the planning phase could also be delegated to a 'lead evaluator' and not all evaluators need to be present."(Rd. 2*). Further, evaluators must physically attend the exercise. This may hold for field-exercises and command-post-exercises without any limitation, but there might also be, e.g., discussion-based exercises, that may be created in a digital environment, where virtual participation can be imagined. He/she further argued, "*in that case if the exercise is partly digital the evaluators need not be present.*" (Rd. 2). Another respondent elaborates on that, "*In exercises with physical participation (majority ones), the physical participation of the evaluators is also a prerequisite. Exception can be computer assisted exercises that players are from their offices and the exercise occurs online via an assisting tool. This relatively new type of exercise, very effective for multi-organizational and national exercises, should not be neglected." (Rd.*

[2] "Availability" is taken as an indicator of *temporal* availability. However, in the analysis and interpretation stage, the project team recognized that it may also be understood as an indicator for being accesible by phone, radio communication or else. This different interpretation of the concept must be taken into consideration.

2). Here, the mode of attendance seems to correlate with the mode of attendance of exercise participants. Further, one shall critically reflect on that even if the evaluator is attending the exercise on site, it may not be possible for him or her to be present at each site, which must be taken into account, "*one could also question if it is actually possible for evaluators to be physically present at all exercise locations. Here, the objectives and focus of the evaluations should be leading to prioritize involvement." (Rd. 2).*

Compliance with formal and informal behavioral rules: Evaluators must understand how to properly act during the exercise. They have to understand and comply with the rules of engagement and safety measures at the exercise site(s). In the second round, 7 agreed (n=7) upon that. One respondent highlighted that "*evaluators must be a plus and not a concern of exercise organization."* (Rd. 2). Despite agreement, one critically commented that the statement was rather vague, "*sounds great but is rather vague: what is proper behavior, of course someone needs to behave properly, but what does it mean*." (Rd. 2) Thus, it is from high importance to elaborate on the expected role behaviors and the general rules of engagement in greater detail during the exercise planning (Code of Conduct), which shall then become a key part in the training/briefing of the evaluators. Beyond compliance with formal rules, intercultural competence and knowledge of the local culture are valued competencies. Although not being a knock-out-criterion, evaluators should be aware of acting within a specific context, i.e., in foreign countries and within culturally heterogeneous teams. In round 2, 5 experts (n=7) agreed on that without any limitation (e.g., "*Indeed it is not a prerequisite but it is a competence that should be considered when recruiting*"), while others suggested greater importance. One respondent commented, "*they should indeed comply with this, especially the context of international cooperation should be nothing new for them, as they need to be able to act proper in an international environment with different cultures." (Rd. 2*). One expert agreed, but suggested to consider various subcultures as well, "*Fully agree, but not only local culture is important, any (sub-) culture (incl. UCPM 'culture') awareness is important." (Rd. 2)*

Comprehension: Evaluators must be able to identify problems and deviance (analytical skills, text comprehension) and understand it within their specific context (contextuality): *"My experience with evaluating and evaluators is that they need to be able to apply a certain level of logic and therefore need analytical skills." (Rd. 1).* In the second round, 7 agreed upon that (n=7). 1 emphasized that this may lie at the heart of an evaluation, "*evaluation implies comprehension what you see and write it down in an analytical way" (Rd. 1*). 2 suggested a more comprehensive wording not being limited to identify problems and deviances, e.g., *"Evaluators should not only identify (this is more related to observation) but also be able to analyse (not only deviances). Context (in its broadest sense) is very important and should be taken into account in the analysis phase that leads to a conclusion (judgement)." (Rd. 1)*

Further, ability to assess was also rated with rather high priority, but on lower consent. This may again reflect the variety of evaluator roles and focus. However, in round number 2, just 2 agreed on that (n=8). It is suggested that ability to assess is of great importance, e.g., "*I believe that this skill should be given a much higher priority. What is written here creates a confusion: An Evaluator is NOT a Note Taker! He can do both, but if we are talking about Evaluators, then the ability assess [sic!] what she / he observes, and whether this is relevant to her / his task is perhaps the most fundamental aspect of the process." (Rd. 2*). While one argued, that this is crucial for each evaluator *("in all positions assessment is important to reveal unexpected views of the exercise")* others understand assessment as a particular task assigned to particular roles and functions, e.g., "*I would suggest various phases as mentioned in the text so you have various roles/functions (that can be separate or in one person i.e. the evaluator): observers that collect observable data, analysts (who analyse data) and authors." (Rd. 2*) Indeed, mainly the lead evaluator and some evaluators appointed by the lead and the exercise management are analysing the data collected."

Organizational talented: Even though evaluators should feature organizational skills, staff management was rated with lower priority. These skills may add value, but are not equally required for each evaluator, e.g., required for lead evaluators. 5 agreed upon that, 1 rather agreed (n=6). 1 emphasized, "*Yes as evaluators are added value and not concern for exercise organisers and may need to serve in other roles if needed." (Rd. 2)*

Language skills: Clearly, evaluators must be able to communicate and understand and comprehend communication between the exercise participants.

High level of English proficiency: On average, respondents declare an English proficiency level between B2 and C1 according to the CEFR as a minimum requirement, listening and reading skills being rated somewhat higher than speaking and writing skills. While, in round number 2, 5 fully agreed, 1 commented on the need for communicative skills (see below), which is more about bringing language in action, so to say, than formal language skills, and one questioned, *"I wonder why passive skills are more important that active skills....an evaluator should not only be able to collect but his/her most important role is to share the evaluation outcomes and therefor active skills are more important. So in order to fulfill its role active skills and proficiency is more important" (Rd. 1).*

English is the working language of the UCPM exercises. Further, English is required during actual disasters and international civil protection missions. One respondent in round 1 clearly prioritizes English proficiency compared with local language proficiency and highlights, that not just evaluators, but also local actors need to communicate in an international language: *"local people must be encouraged to speak, write and understand international languages and especially English as FSX are international and in reality, many foreigners will be present during actual disasters."* (Rd.1)

Further, English proficiency requirements also vary along roles and tasks to be conducted. Lead evaluators, for example, are required a higher level of English proficiency. In round 2, 5 agreed on that (n=5).

Finally, language requirements also vary along with exercise types. E.g., for an observation and evaluation of technical skills, even a lower level of B1 might be sufficient, while the evaluation of communication during a high level tabletop exercise (TTX) may require at least C1, one respondent illustrates. This is confirmed by another respondent, who argues, that "*for the tabletop (discussion-based exercise), the level of language (speaking and listening) should be higher, as by definition everything is discussion-based*." (Rd. 1). In round 2, 4 of 5 agreed on that, while one was wondering about that, "*can you provide reasoning behind this? I can imagine that this is due to the possibility to interact and clarify."* (Rd. 2)

Team members with local language skills: The required level of proficiency in the local language, on the other hand, is rated rather low. If declared as a requirement, basic language skills were declared sufficient with slightly higher values for speaking and listening skills. In round 2, 5 of 5 agreed on that, one putting strong "*emphasize on 'recommendation' not as necessity" (Rd. 2).* However, it may clearly add value to exercise evaluation, as several respondents highlight, and it is highly recommended to have an evaluator team, with some being able to understand local language: "*the inclusion within the team of evaluators of persons speaking, writing and understanding the local language is very useful as all types of exercises and evaluators' tasks, however, cannot be a prerequisite for selecting evaluators." (Rd.1*) Another one argues, "*It is always good to have "local" evaluators who know the local language and English at a good level, which makes the job somewhat easier, but if the evaluation coordinator knows both languages and the exercises are in the local language, there is no problem with later translation" (Rd. 1*). However, others argued that the evaluation team may also be provided with assistance for understanding local language, e.g., by providing specific functions such as "*an interpreter" (Rd.* 2) or "*an Evaluator Liaison Officer, similar to the one(s) used for EUCPT" (Rd. 2)* or technical assistive solutions, such as a "*translation software" (Rd. 2).*

Further, the priority of local language proficiency also depends on the evaluation objects and objectives. In international exercises, English language proficiency is considered sufficient for an evaluator focusing on the communication and interactions between the European Union and member States and the local emergency management agency (LEMA), for example. However, "*If the focus is on the LEMA then local language skills are required (but not for all evaluators, only the evaluator focusing on that part*)." (Rd. 1). That "*specific tasks and objects may require local language skills*" was supported in Rd. 2 as well, while another one critically argued that the capability to understand the "smooth" (or not) functioning of LEMA is not necessarily linked to local language skills, "*even if the knowledge of local language is absent an evaluator must be capable to understand smooth of LEMA functioning or not." (Rd. 2).*

Practical language: Evaluators observe and analyse the communication, actions, etc. of the members of the emergency management system. Thus, they should also understand formal and informal language used within emergency management and the civil protection community, such as formal terminology, commonly used abbreviations, and colloquial language.

Affiliation & Experience vs. Alienness: Interestingly, the need to be affiliated to a first responder organization and previous experience with exercises or exercise evaluation was rated with rather low priority, with relatively high dissent for affiliation to a first responder organization (Rd. 1). This contradicted our expectations, but to some degree is in line with recommendations to favor external evaluators, who do not have close ties to the evaluation objects and are not confronted with conflicts of interests.

In round 2, it became even clearer that the relevance of experience and affiliation is quite controversial. 4 of 5 agreed that affiliation and practical experience may add value but is not necessarily required. One position highlights the benefit of being alien, "*Many times you need the different sight of people with no experience in exercises evaluation to see things that we can not" (Rd. 2),* whereas the other position is characterized by a reluctance against the lack of experience, "*without any previous experience one cannot evaluate properly how a rescue exercise/operation is going" (Rd. 2).* The same pattern occurred regarding the perceived benefits of being alien to the organizations and system. 4 agreed (n= 5), with one respondent strongly supported it, e.g., saying that "*Fully agree. Need to have a "clear" picture that most people closely engaged to an organisation are used to and do not notice." (Rd. 2*). Others pointed to the fact that being alien must not mean a lack of *understanding,* e.g., "*Certainly, but you do need to know the relevant procedures!", and both alienness and familiarity are ambiguous features, "could be an added value but is a disadvantage at other parts of the evaluation" (Rd. 2).*

While this is highly informed by epistemological and methodological stances, the contradiction becomes even stronger, when considering practical problems that may occur. 4 out of 6 agreed on the statement, that external evaluators, who are affiliated to the same *kind of organization,* should be preferred to internal evaluators. However, 2 respondents strongly contradicted this statement, one underpinning the benefit regarding the ownership of findings, "*This is a debatable item. As we are not talking about a certification (INSARAG-type), the non-involvement of the conducting organisation (in particular) in the evalution process might result in findings which are not 'owned' by the people who should care most about them. DG-ECHO has been particularly sensitive about this aspect in its exercises, and rightly so (I feel)"* (*Rd. 1),* one considering an effective feedback loop, "*addition I really question if it is necessary to have (solely) have [sic!] an external evaluator if the purpose of an exercise is learning. Internal evaluators with a larger distance to the evaluation object might be even better as the knowledge and insights gained remain within the team." (Rd. 1).* Again, it seems that a certain attitude and soft skills (objectivity, critical distance) is far more important than formal characteristics. Further, it becomes clear, that a good evaluation must not just take into account a well performed evaluation in terms of theory and methodology, but also must consider, what is required to put the evaluation in

practice. The "best" evaluation, so to say, would be meaningless, if it is neglected and not paid attention to its finding.

This tension is partly solved, when considering the competence area of professional knowledge, where the need for *understanding* was expressed in contrast to mere professional *expertise*. Finally, having mixed teams in terms of practical experience and affiliation may be a good clue, thus including evaluators that are familiar with the system and organizational processes and structures, and evaluators that are alien to it. In round 2, this was again raised by an expert, "*a mix should be preferred." (Rd. 2).*

Cognitive maturity: The minimum required formal education was rated between uppers secondary school and Bachelor's degree. One respondent remarks, "*It's not rocket science, but necessary with at least upper secondary to be able to understand the procedure" (Rd. 1).* Evaluators must be able to understand various concepts, organizational structures, and key competencies such as coherent writing, etc. and formal education may indicate these.

However, respondents are quite reluctant against a too rigid declaration of a minimum education. Again, soft skills and attitudes are mentioned as far more important than formal certificates or alike. E.g., one respondent highlights this issue: "*An academic degree offers many skills that are useful for an evaluator BUT many useful skills are not learned in school, which is why I would not set a minimum education a status. In any case, relevant work (or volunteer) experience is far more important than formal education*." (Rd. 1). Here it becomes clear that formal certificates *may* be an indicator for skills and competencies, but neither a guarantee nor a logic precondition. Thus, the requirement may rather be termed cognitive maturity than formal level of education. One respondent elaborates on that: "*More important is the thinking level than the actual formal degree. A thinking the level [sic!] of preferably bachelor is good." (Rd. 1*). In a similar vein, another respondent argues in favor of academic education, but also emphasizes the need for critical reflection, "*my experience with evaluating and evaluators is that they need to be able to apply a certain level of logic and there for [sic!] need analytical skills. Evaluation is comparable to research and this is mostly provided a bachelor/master level. Maybe it is better to indicate competences instead of education status. Because some education is less suitable."* (*Rd. 1*) Again, some diversity in the team composition is favorable. Teams shall be mixed in terms of expertise and the occupational background. Besides academics, also "*people of practice*", one respondent writes, shall be included. Another one elaborates on this argument: "*Obviously, a degree provides a certain status especially for written skills, ability for data elaboration etc., however it depends on the specific role of the evaluator and the capability he/she is appointed to evaluate. Depending on the role and position of the evaluator, his/her practical experience can be more valuable, with some minimum of skills. Different backgrounds and profiles, of a practitioner or of a manager or scientist can be required." (Rd. 1*) In round 2, 5 agreed (n=5), one of them further arguing, "*I think this should not be pretested as an independent variable. As you already mention practical experience might be equivalent. Maybe other competencies such as language skills, analytical skills, observation skills, etc. are important and this is not only taught in formal education." (Rd. 2).*

Ability to navigate: Spatial skills (orientation, analog navigation, navigation system) were rated with moderate to low priority. They add value but are not to be considered a kick-off criterion. However, evaluators must be able to understand exercise plans and maps and to navigate through the scenery/sceneries to handle the situation.

Orientation towards action problems: There was dissent on the evaluators attitude towards action problems (problem-oriented, solution-oriented, creativity). While the mere evaluation not necessarily depends on these features, they may support evaluators in the evaluation situation. Creativity, for example, provides him/her with the ability to find ad-hoc solutions for unforeseen challenges.

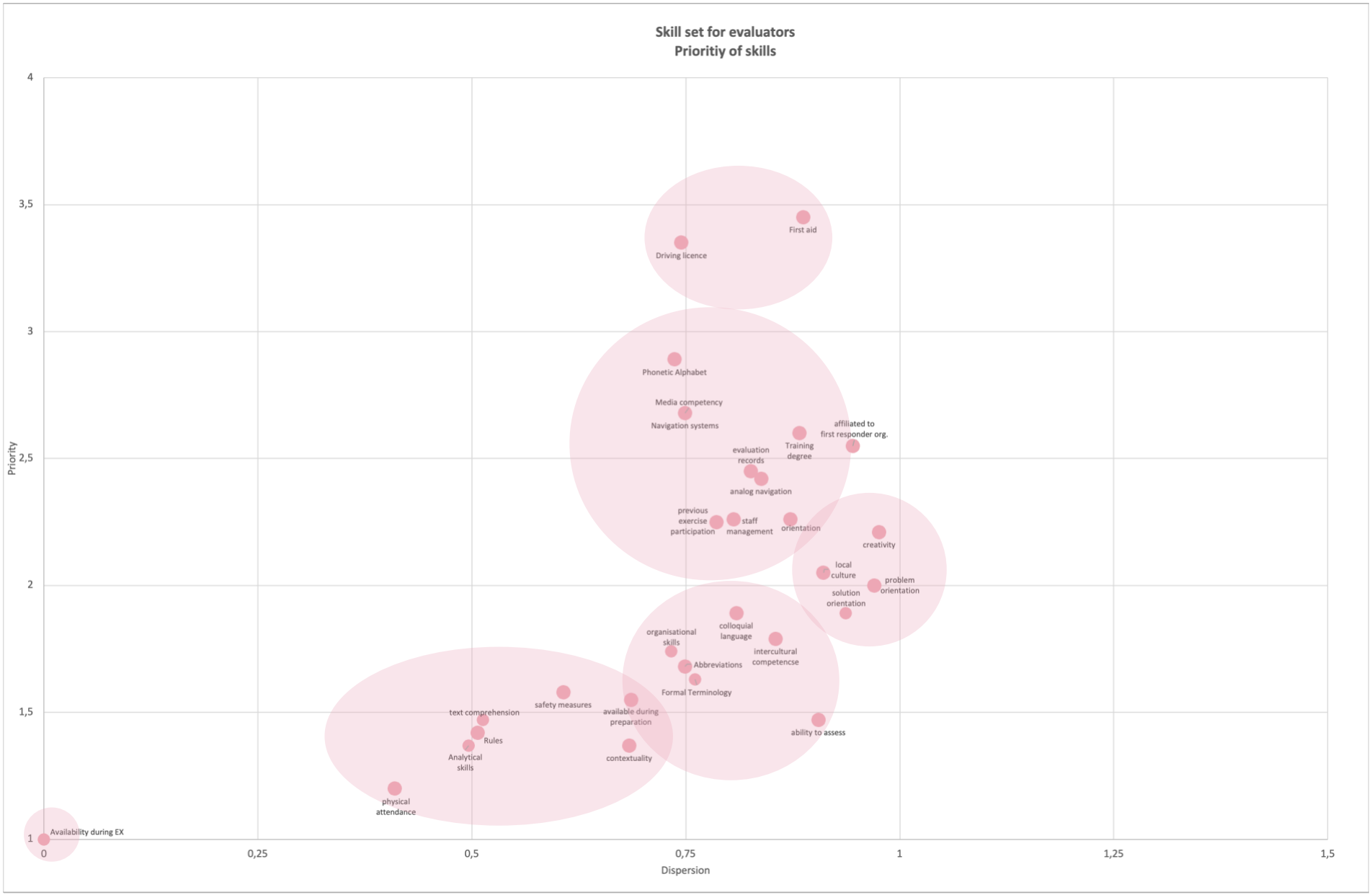

Figure 3: Prioritization of basic skills and competencies

3.2 Professional competencies

Interestingly, professional knowledge on diverse subject matters was rated with moderate priority and slight differences in the level of consent. The reults were aggregated and interpreted as follows:

Professional knowledge on the evaluation objects: For one thing, the broad variety of exercises comes to the fore. Which professional knowledge areas matter the most, is a matter of the evaluation objects, objectives, and exercise goals: *"Which compentence is needed depends on the objective of the exercise and compentences related to the objectives should have high priority, but this could be spread by priority on different evaluators. F.e. the system is important for an objective related to coordination, while the technics are more important for an objective related to the performance of a specific module" (Rd. 1).* Another argues: *"The professional compentences depend on the exercise and moreover evaluation aims. If the focus of the evaluation is technical then the evaluator might require some more background in this and thus it is a high priority. If the focus is on the system than the systems perspective is more important" (Rd. 1)*. Again and again, the respondents emphasized that especially regarding the required areas of expertise, there is no answer that fits exercise evaluation in general: *"This question is impossible to answer without linking to the ToR [Task or Role] of the evaluator and the exercise objectives. E.g; if USAR [urban search and rescue] tactics is an objective it requires a different skill set than evaluating the Information management in the work of a strategic staff" (Rd. 1).* There simply is no general requirement for an evaluator to cover knowledge in particular areas, but rather, evaluators with various areas of expertise shall be properly selected in line with the planned exercise and its particular goals at hand: *"THE evaluator does not exist and these criteria can be used for selection and then the priority is given, now they are basically equally (un)important" (Rd. 1).*

Knowledge on the evaluation objects (what is to be evaluated) is of high importance, one respondent argues in line with the others, while, in this regard, the type of exercise and exercise context (scenario being trained), matters less: "*The knowledge of the evaluators depends on what should be assessed and what should be strongly emphasized during the training, then the type of exercises does not matter. It's good when the evaluators have an experience to be evaluators then the knowledge about the specific scenario is in the second place" (Rd. 1*). This was also supported in round 2 (5 of 5 agreed).

In round 2, it became even more obvious that *knowledge* on the specific evaluation objects and exercise goals being trained is needed. 4 of 5 agreed on that; one highlighted the need to differentiate between *expertise/professional knowledge* and *understanding: "At least they should be able to gain this experience, but the level depends...be careful that this does not contradict earlier statements e.g. an evaluator does not have to be an operational commander or so to evaluate them, they rather should understand their role. Thinking about this 'understanding' might be a better term than expertise." (Rd. 2).*
Accordingly, 3 of 5 experts agreed that it is not possible to list the specific areas of expertise prior to the planning of a specific exercise. However, the other 2 remarked, that, again, *expertise* shall be changed with *understanding,* and that a broad understanding or overview of disaster management and civil protection adds much value, thus might be considered a universal requirement, because this also relates to what was earlier considered "contextuality": "*Broad experience offers significant added value as many things are connected in civil protection and many times professionals miss the link (i.e. disabled people and civil protection, tourism sector engagement etc). So one must see the broader objectives and links of an exercise and decide on expertise needed." (Rd. 2)*

What is from outstanding importance, is that evaluators must have a solid understanding on their own role and functions in the exercise, and how their evaluation may impact the field of civil protection. 5 of 5 agreed upon that in round 2.

Finally, there was disagreement on the conclusion that expertise is not needed a priori, because sufficient knowledge on the specific areas can be trained beforehand. While, again, 3 of 5 agreed upon that statement, 2 critically remarked that it depends on the time horizon of the training program, or questioning the possibility at all: *"I do not see how this is possible, an evaluator should know what he is asked for so he/she can be prepared." (Rd. 2)*

Going native (to some degree): In line with the comments above, the role of having "material" knowledge or specific "expertise" was, to some degree, a controversial issue. One respondent argued that being alien to a system, or procedures may add fundamental value: "*Some times weaknesses are revealed by people not so experienced in a topic, so not used to the procedures, and thus able (if they have some broader perspective of the situation) the issues that need improvement. It is like external consultants in a business (the independent eye)"* (Rd. 1). This is in line with emphasize on independence and critical distance mentioned below (personal competencies) and the relatively low priority of being affiliated to first responder organizations and exercise experiences (both evaluation records and previous participation).

Further, 4 of 6 have agreed on the statement in round 2. However, going native must not be mixed up with lacking knowledge at all. One expert critically remarked that "*lacking relevant knowledge could be a problem for the whole evaluator team, which could outweigh the added positive effect of the possibility to find a blind spot" (Rd. 2).* Another one agreed but pointed out that *"it is also good to operate in teams with various backgrounds, experiences, etc."(Rd. 2).* And elsewhere, another expert remarked that *"the evaluator must be knowledgeable in at least something related to disaster management" (Rd. 2).*

More and more, it seems that being evaluator requires a certain attitude, role understanding and systemic thinking, rather than a bundle of expertise on subject matters: *"It is important that the evaluator understand the scope of the exercise and the importance that his role have on its outcome, and thus to perform his tasks with professionalism and seriousness." (Rd. 1).* In round 2, also 5 out of 5 agreed that being evaluator seems to be much more about softs skills and key competence, rather than a specific "list" of hard skills.

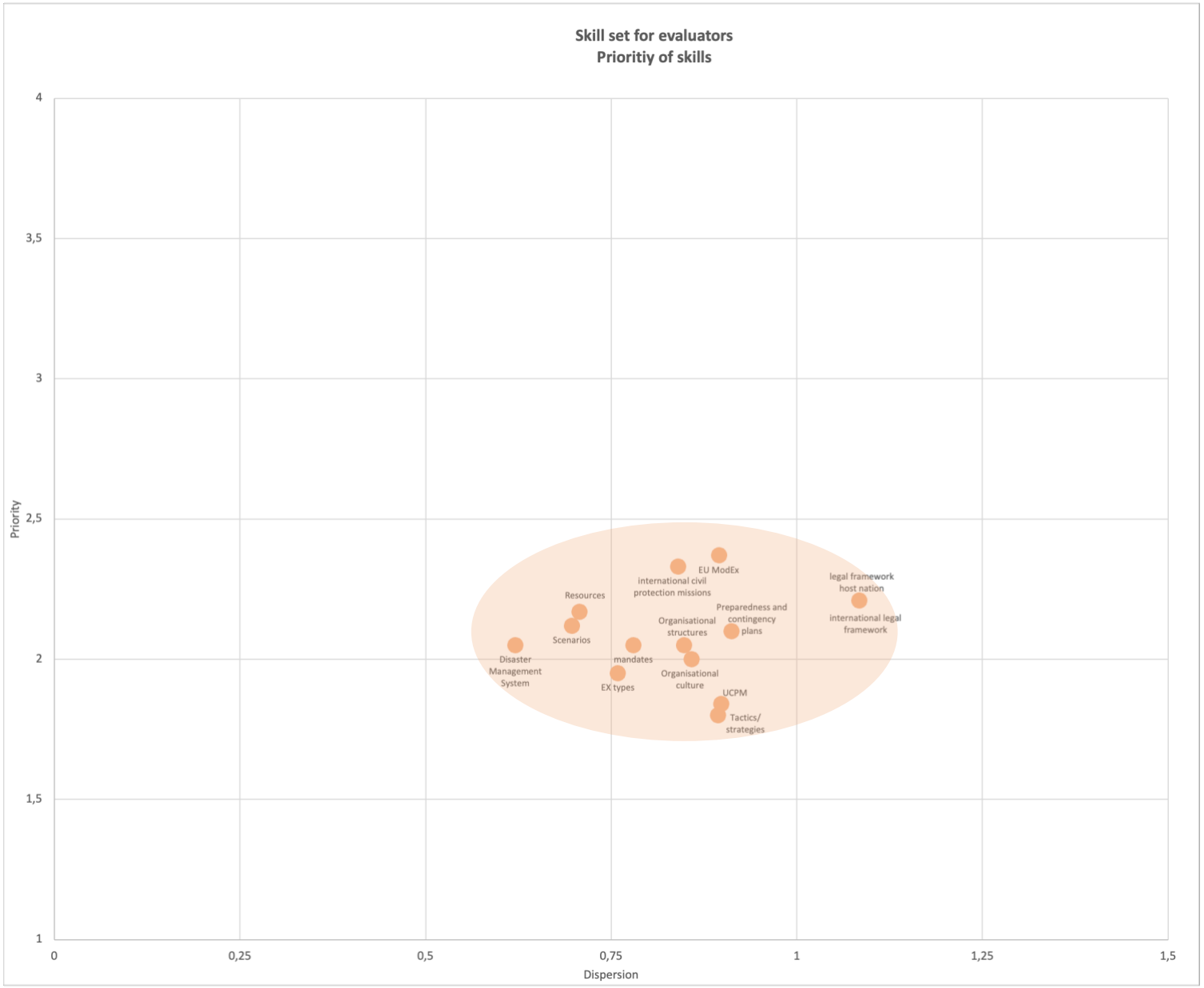

Figure 4: Prioritization of professional knowledge

3.3 Methodological competencies

Methodological competencies were, on average, rated as follows:

- **High priority/consent**: Respondents rather agreed that the ability to conduct observations and to use computers and forms and checklists are of high priority.
- **Moderate-high priority/neither nor**: The ability to use messenger services, evaluation apps, smartphones, cameras and VCT were rated with moderate to high priority. The same applies to data gathering methods such as interviews, document analysis, surveys, and data processing and analytical methods. Meta-competencies such as technical literacy, and the ability to select the proper methods were also rated with moderate to high priority. The ability to hold After-Action-Reviews and Debriefs was also rated with rather high priority, while workshops in general were rated with moderate priority and low consent.
- **Moderate-low priority/neither nor**: General methodological competencies in the area of qualitative or quantitative research methods, as well as the opportunity to use radio communication was rated with lower priority.

This may be aggregated and interpreted as follows:

General methodological rigor: Methods shall be applied with rigor. Methodological skills and competence seem to be rather important. One respondent emphasized: *"Complementarity of methods and tools is the key to a successful evaluation." (Rd. 1).* Another respondent criticizes: "*Many exercises don't progress beyond basic observations and some scribbled notes. Proper data collection, management and analysis is often lacking." (Rd. 1).* In round 2, 5 of 5 agreed, with one pointing out that methodological rigor is a cross-cutting issue, "*However this also influences the previously identified competences. It is all connected." (Rd. 2)*

Basic tools and methods: Evaluators must generally be able to apply basic tools, such as computers and forms and checklists, and particular methods, i.e. observation. Beyond that, hardly any methodological competence may be considered a universal requirement for evaluators *in general*. In round 2, 4 of 5 agreed, while one critically questioned, "*What about other methods such as interviews, surveying, focus-groups, group-sessions are not important as methods for an evaluator?" (Rd. 2)* This ambiguity can be easily resolved, when considering that the variation in the evaluator's roles and the exercise types, scale and objective. These require the application of certain tools and methods, while observation and computers and checklists may be perceived as standard repository of evaluators that are important for each role and task.

Methodological meta-competence for leads and organizers: Firstly, which tools and methods are to be prioritized in the selection of proper evaluators varies along with exercise characteristics such as type, scale and objectives meaning that not each skill counts as skill in each context and vice versa. Here, the priority of methodological meta-competence comes to the fore, i.e., the ability to select the proper tools and methods, as one expert emphasizes, *"What is to be evaluated must be observable. It is important that the evaluator can decide how the evaluation may best observe the evaluation subjects to meet the goal." (Rd. 1*) In round 2, 4 of 6 agreed on that. One highlighted that this is to some degree at odds with the endeavour to develop a standardized methodology, *"If we go towards a unified methodology of evaluation it is not necessary, however, ability to judge methodology is important for improving it." (Rd. 1)*
However, even meta-competency was not rated on high priority with consent, but moderate to high and lower consent. Not each evaluator is equally required to feature methodological meta-competence. It may be the most required for those, who plan and organize the evaluation approach.

This is supported by an expert in round 2, saying that *"This is certainly not relevant for all members of an Evaluation Team, but mostly for the Planners / Designers of the Evaluation process" (Rd. 2).*

Task-related skills: Secondly, it varies along with the role of the evaluator and the assigned tasks. In other words, evaluators must be capable of those methods and tools, that they are expected to apply during the exercise evaluation. One respondent, for example, emphasized that it makes a huge difference if the evaluator has merely to observe practical work or also to conduct interviews and analyses. Another referred to the difference between particular mandates and evaluator in the field, "*the methods are all very important for an exercise evaluation organizer/lead evaluator/writer of an evaluation report, but this does not imply that an evaluator needs to know this." (Rd. 1)* This is confirmed by a further respondent, who emphasizes that *"some of the competencies, e.g., data analysis is of course good to have, but they are the most important for the Lead evaluator and its direct team that will do the data analysis and processing based on research methods, and not for every evaluator in the field." (Rd. 1).* In round 2, 4 of 5 agreed, one of them adding that, therefor, training is necessary, and one expert criticizing that this is a too vague description.

Methodological competence portfolio to find the best match: There are only few competencies that are to be considered "universal" requirements for evaluators. Beyond that, evaluators should provide a comprehensive portfolio in terms of expertise (substantial knowledge), methodological competence, social and personal competence, to support proper selection and recruitment of evaluators, but must not be taken as kick-off-criterion: *"They can be used as a indicators/variables for selecting evaluators fit for the job but in general they are equally important for performing evaluations... but not for specific evaluations." (Rd. 1).* 5 of 6 agreed on that in the 2nd round. However, the issue has been raised that there is a need to define some standard repository methods and tools, *"the UCPM might support the evaluators with common methods are preferred methods in EU exercises" (Rd. 2*). Once such a standard repository is defined, there is no longer a need to provide a comprehensive portfolio. Again, it was argued that "*if we go towards a unified methodology of evaluation, it is not necessary, however, ability to judge methodology is important for improving it" (Rd. 2).* Finally, one expert remarked that this may be, again, "*not relevant for all members of an Evaluation Team, but mostly for the Planners / Designers of the Evaluation process" (Rd. 2).*

Preconditions vs. training agenda: Further, characteristics of the tools and methods should be taken into consideration, i.e., the easiness of learning how to apply them. Again, it is more about soft skills and meta-competencies than hard skills. One highlights, *"if one doesn't have already the capacity to use them (e.g., software supporting evaluation tasks), they can be learned easily, and I don't think this should exclude a competent evaluator from performing his/her role if not very knowledgeable on the tools." (Rd. 1).* While any implicit knowledge, such as digital literacy, may hardly be trained prior to a given exercise, certain tools can easily be introduced and learned during exercise briefings, if the evaluators feature key competencies. In round 2, 4 of 5 agreed on that, and 1 raised the idea, that "*an introduction course for evaluators should be designed (for each exercise or in general*)" (Rd. 2).

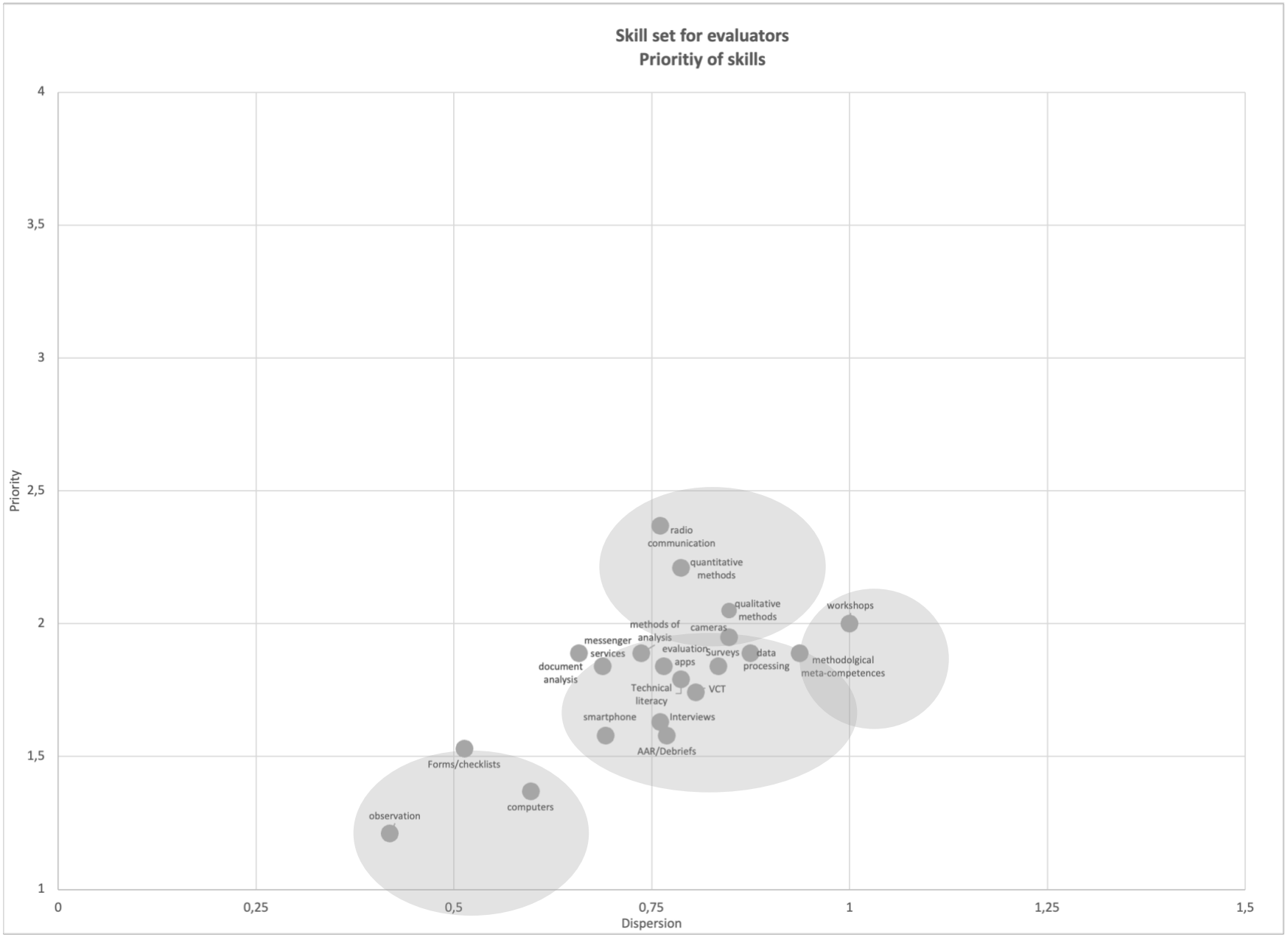

Figure 5: Prioritization of methodological competencies

3.4 Social competencies

Social competencies, i.e., the ability for successful social interactions, were rated as follows:

- **High priority/consent**: Communicativeness, cooperativeness, appreciative attitude, adaptability, tolerance, transparency, and the ability to give constructive feedback were rated with high priority and rather high consent.
- **Moderate to high priority/neither nor**: The openness to criticism, thus the capacity to receive feedback, was rated with moderate to high priority.
- **Moderate priority/low consent**: Rhetorical skills, empathy, consensus-oriented attitude, and conflict ability were rated with moderate priority and low consent.
- **Moderate-low priority/consent**: Assertiveness and being reserved was rated with moderate to low priority on a higher level of consent.
- **Low priority-not required/neither nor-dissent**: Personal traits such as introversion or extroversion were rated with very low priority to not required. The same applies to competitiveness, leadership competency, subordination, and willingness to compromise.

The findings may be aggregated and interpreted as follows:

Communicativeness: Evaluators must be communicative and transparent. They need to express their roles and functions within the setting, and to communicate results and lessons learned to the clients in a way that can be easily understood: *"Social skills are essential for an evaluator fulfilling its role and underline that he/she is there not as a protagonist but as a facilitator of change/improvement." (Rd. 1).* Each agreed on that (n=5). Regarding the former, one added that "*evaluators are there to be seen as supporters and not judges and that must be clearly communicated" (Rd. 2).* Regarding the latter, one added that, "*making findings understood but all participants are important especially in cases of non-professionals participation to FSX" (Rd. 2).* Obviously, evaluators are often confronted with reluctance. One expert argued, "*one could also question why an evaluator needs to 'defend' his/her role. This has to do with a more fundamental 'problem' in the way evaluations are used and also perceived. Here there is not only a role for the evaluator but for the whole civil protection community to demonstrate the added value of evaluations and remove any fears (but we need to know where these 'fears' come from...)" (Rd. 2*). This may also go as far as aggressive attitudes towards the evaluation teams. One expert problematizes experiences of "*aggressive attitude towards evaluators and disagreement with their findings*." (Rd. 2). Such experiences may also underlie the bifurcation regarding the attitude towards conflicts and problems (consent-oriented, conflict-oriented).

Team spirit: Further, evaluators usually work within a team. Against this background, respondents agree that that evaluators must be team players (cooperativeness), and adapt to teams instead of doing his/her own business. Further, it is important that evaluators signal that they work towards the same goal as participants, i.e. to assure quality within civil protection, thus, to create the image of evaluators and exercise participants being *one team*. Evaluators, as mentioned above, need to depict themselves as "*facilitator of change/improvement*", not as enemies.

Respectfulness: Indeed, evaluators are in the difficult position to illuminate gaps, weaknesses, failures, and similar. This may be a sensitive issue. Thus, evaluators must feature a respectful attitude towards his/her clients. It is essential that they are able to give feedback in a constructive manner and have a tolerant and appreciative attitude. 4 of 5 agreed on that, while one critically remarked that this "*Depends on the evaluation type e.g., aprreciative evaluation might be an approach but also other approaches can be chosen" (Rd. 2*). Another respondent warned against relativization of findings,

"*Respect, yet independent and not afraid to propose improvements aftehaving explain that we are on the same boat" (Rd. 2).*

Task-related skills and competencies: Beyond that, consent is rather low, also for those competencies, with moderate priority. Generally speaking, social skills and competence are essential for a successful evaluation, but – likewise to the other competence areas – there is no such a thing as a closed list of required competencies. Unless social skills are not put into practice, it is hardly to define a clear list of requirements: "*People are different so you do no[t] a priori which set of social skills will create the maximum benefit and it is better that way since it is closer to reality." (Rd. 1).*

Thus, the priority of the various competencies highly depends on the characteristics of the evaluation. The roles and tasks of evaluators are pivotal, as one respondent mentions: "*Here it depends on the role of an evaluator: is the evaluator only observing and writing down observations or is it a person having interaction with the persons trained?" (Rd. 1)*. In a similar vein, another respondent argues: "*this also depends on the type of evaluation that is performed. As you know there are many types of evaluation varying from measuring, analyzing, describing, etc. and various ways of providing feedback (verbal, written) this requires different competencies to be important. For example, an evaluator that needs to provide a verbal hot wash in front of a large group might not be introvert and should have good communicative skills and a good level of proficiency in, e.g., English, whilst if the focus is on a written evaluation different competencies are needed. Also, if the evaluation is more of an appreciative inquiry, it requires a different set as if the evaluation is aimed a consensus or independence." (Rd. 1)*

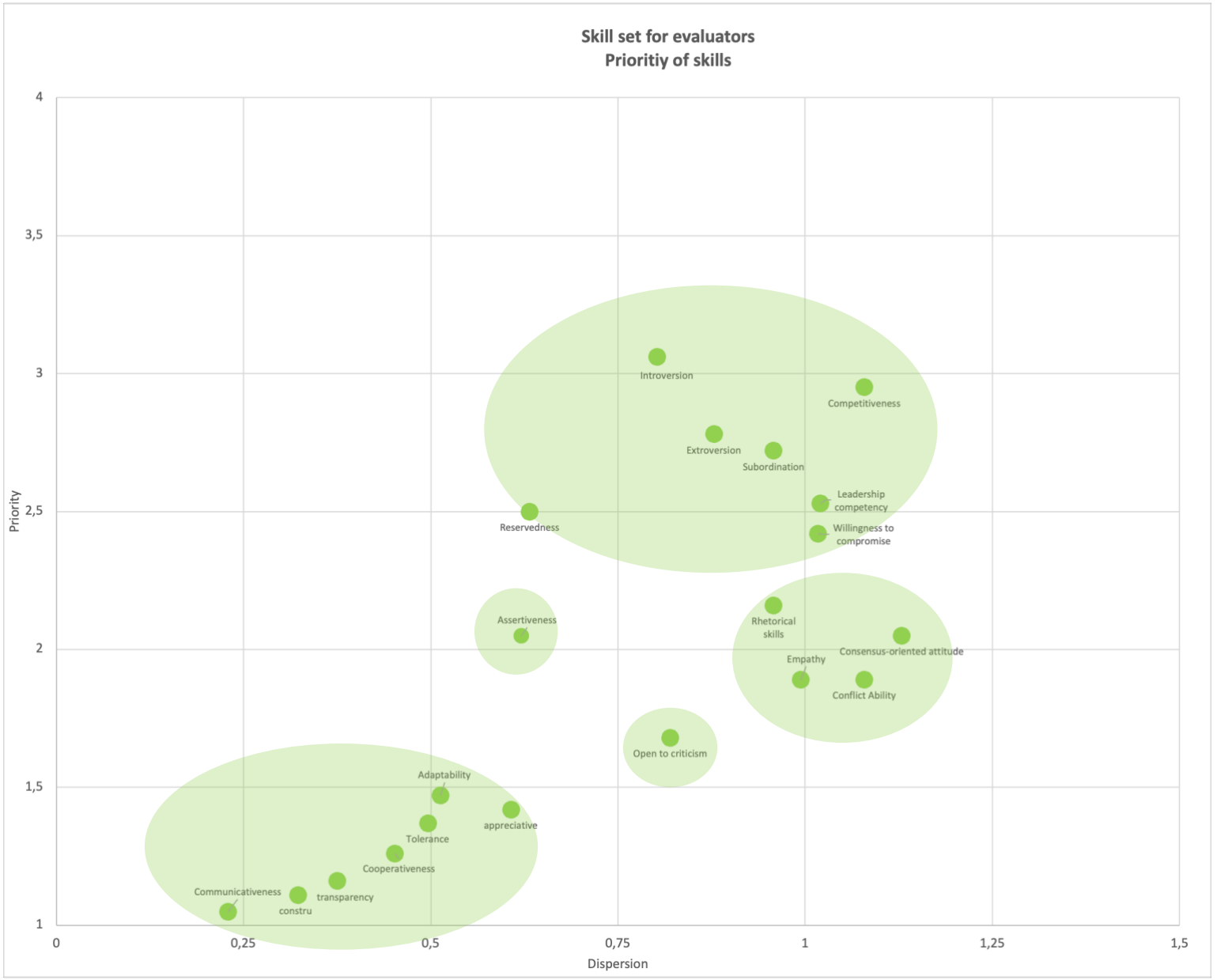

Figure 6: Prioritization of social competencies

3.5 Personal competencies

- **High priority/consent**: Referring to personal competence, the following characteristics of an evaluator seem to be out of question: objectivity, reliability, integrity/trustworthiness and flexibility. They were ranked with high priority and with consent.
- **Moderate-high priority/neither nor**: Further, a bundle of personal competencies was, in average, rated in the range from moderate to high priority: "Learning capacity" and "critical self-assessment", "Proactivity", "Decisiveness", "Resilience", "Motivation", "self-discipline", "accuracy", "pragmatism", and a "questioning/critical" and "holistic" perspective.
- **Moderate-high priority/dissent**: "Passivity" was also rated with medium to high priority, but with relatively strong dissent.
- **Moderate priority/neither nor**: "Pedantic attitude", "rapidity", and "self-determined-working style" were rated with moderate priority and neither consent nor dissent. "Idealism" was also rated with moderate priority, but lower scattering.
- **Low-moderate priority/dissent**: "Prudency", and "externally determined working style" is rated between moderate and low priority with rather high deviation.
- **Not required-low priority/neither nor**: Finally, impulsiveness was rated in the range of low priority to not required.

The findings are aggregated and interpreted as follows:

Independency and ability to keep a critical distance: Evaluators must be independent observers of the exercise. "Objectivity", i.e., the ability to keep a critical distance and apply an objective perspective on the exercise and evaluation objects, was rated by each of the respondents with high priority. One respondent argued: "Many are to assertive, i.e., imposing their own experience and views and using that to reflect on the performance, instead of an open mind and analytical eye." Further, "integrity and trustworthiness" was rated with high priority and consent. This is an indicator for the level of trust that evaluators do the best of their ability to make their assessments free from influences by third parties. This is also confirmed by prioritizing a "questioning" attitude and "holistic" perspective compared to a "pedantic" attitude. Evaluators should be able to keep a critical distance from their evaluation objects and keep in mind the big picture, rather than a too strong focus on details. This is also in line with the rather low priority of the affiliation to first responder organizations described above and the strong need for external evaluators described in the handbook (Swedish Civil Contingencies Agency 2011), which may lead to a too narrow focus or self-identification with plans, processes or structures.

In round 2, 4 of 5 agreed upon that. However, one critically remarked, "*Is independence the same as objective? Would objectivity, integrity and trustworthiness not be more important than independence, in particular in 'learning environments'." (Rd. 2).* Elsewhere he or she again questioned*, "is it needed to be fully independent and is it actually possible within a UCPM context (and being experienced in this field), we can rather be clear about independence that suggest some kind of 'fake' independency." (Rd. 2).* Thus, "independency" may be misleading in its semantic meaning. Here, it is understood as the ability to observe and judge free from external pressure.

Adaptability: Evaluators must be able to handle complex exercise situations. Experts agree that evaluators must be reliable and responsible, i.e., organizers and other actors can rely on the completion of the evaluation tasks, and flexible, i.e., be able to adapt to changing situations and context. This reflects both the variety of exercise contexts, type, objects and goals of evaluations, and the complexity and dynamics within any given exercise (e.g., full-scale exercises). Further, pragmatism

was prioritized compared to idealism. Complex and high-dynamic environments, such as civil protection exercises, it is crucial to keep the right balance between the principles and ideals of evaluation plan and its implementation. This means that one should assure high quality, but still be able to lower one's sight in order to remain capable of acting and completing one's tasks. In round 2, 4 of 5 agreed on that. One added, that adaptability is also needed in terms of the evaluation plans and the own conducts, e.g., when "*reasonable requests*" (Rd. 2) are made by exercise participants. Elsewhere, another expert emphasized in a similar vein, "*the evaluation is built around various pillars such as goals/purposes, requests/questions and these are agreed beforehand. Adapting them during the exercise might influence the setup/design and require addition work or creates a heavier workload that was not foreseen within the original assignment. All parties should then discuss this change and agree upon and discuss the consequences e.g. more time / capacity needed. If it is not possible to mitigate the risks of being flexible one should not proceed as it might effect the quality of the evaluation."(Rd. 2)*

Diligence: Evaluators must be able to conduct their tasks with diligence. For one thing, referring to attitude, again integrity/trustworthiness and reliability/responsibility not just express the need for independence, but also that they act to the best of their ability when conducting their tasks. For another thing, referring to output, "accuracy" (high-moderate) is more important than rapidity (moderate priority). Further, it seems to be a secondary trait, if actors decide and act with prudency or impulsiveness, but prudency was rated with higher priority than impulsiveness. 4 agreed in round 2 (n=5).

Autonomy: Evaluators must feature "Proactivity" and "Decisiveness" and, with moderate priority, "self-determined working style," and "self-confidence." This refers to the ability to conduct tasks self-responsibly and autonomously. Indeed, "externally determined working-style" was rated with low to moderate priority, but dissent was higher. 4 of 5 agreed, one rather agreed, "*but with the sidenote that if they operate within a team these two competencies should be in balance." (Rd. 2).*

Further, passivity was also rated between moderate and high priority, but with rather high dissent. This may resemble, for one thing, the variety of roles in evaluation teams (e.g., evaluation lead, mere observer), and for another thing, the ambiguous role of evaluators, i.e., being both an observer that does not interfere with the exercise and an independent and autonomous actor, responsible for the conduct of his or her tasks. This was supported in round 2, e.g., "*the evaluator shall not interfere, except if there is a risk of injury/damage" (Rd. 2).*

Capacity to grow: The ability to keep a critical distance and to adapt to changing situations also informs the required attitude towards oneself. "Learning capacity" and "critical self-assessment" were rated rather high and with consent. Thus, evaluators not just contribute to the critical assessment and learning within the civil protection community, but also must be committed to grow themselves. 4 of 5 agreed upon that, one putting special emphasize on the meta-evaluation and the development of evaluation and evaluators themselves, "The focus should also be on meta-evaluation to improve the evaluation and evaluator's quality." (Rd. 2)

Engagement: Finally, evaluators need to keep themselves engaged with exercise evaluation. Exercise evaluation and its complexity and dynamics may be demanding, thus, motivation, self-discipline and the ability to handle stress is rated with moderate to high priority. There was full agreement (n=5). Elsewhere, an expert reported the problem of evaluators leaving the team, "Evaluation team members leaving the team which influence the evaluation capacity and quality. One needs to deal with the situation but it was discussed afterwards with the delivering organisation." (Rd. 2).

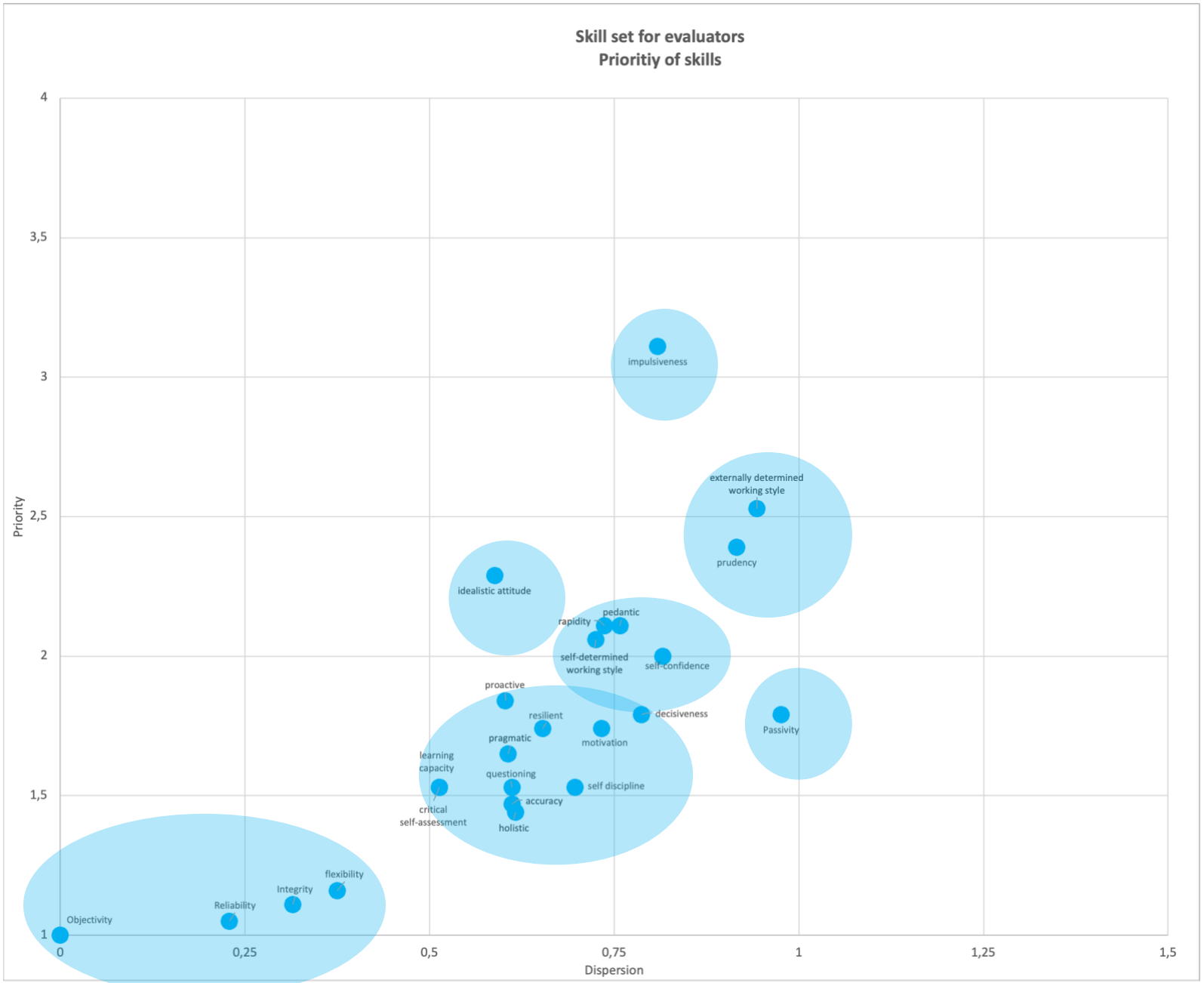

Figure 7: Prioritization of personal competencies

3.6 General issues of exercise evaluation

Moreover, experts were invited to give their opinion on general issues related to exercise evaluation, such as the diversity of evaluator teams, the need for trainings, formalization and standardization.

Figure 8 shows the results for the mean values (dark grey) and the standard deviation (grey). On average, there was rather strong agreement with each item, but still, some items seem to be more controversial than others.

Training of evaluators: There is outstanding agreement on the need for the solid briefings of the evaluators (on average: fully agree, small deviation). Respondents also agree that standardized training programmes may add value, but agreement and consent is lower than for briefings (mean value between fully agree and agree). Agreement on the need for internationally recognized certificates is rather low. There was also full agreement on that during the second round. In sum, the results regarding the training and briefing of evaluators reflect the patterns identified in the prioritization of competence areas as well. Indeed, there is a big variety of exercise types, objects and goals, and context, as well as roles and goals of evaluators. Thus, training for evaluators must be able to keep the balance between flexible adaptation to the cornerstones of specific exercises and some general requirements that hold for each setting. 4 agree on that (n=5). One added that briefing should support contextuality, *"Briefing must also reveal any not obvious connections between the exercise and other fields that make evaluators better understand their role and contribution" (Rd. 2)*. However, one expert critically raised the question about who is going to conduct the trainings, "*I wonder who is going to brief the evaluators i.e. how is this organised and what are they briefed about? Some type of experience in evaluation and training is still needed (or some training aspects such as training interviewing or observing or holding group sessions rather than a general 'evaluators training')." (Rd. 2).*

Team Composition: Referring to team composition, there was rather strong agreement that evaluator teams shall be mixed in terms of occupational background, expertise, experience, and so on. There was full agreement in round 2 (n=5), and one put particular emphasize on heterogeneity regarding expertise and knowledge areas.

However, in some situations, heterogeneity is not required in each dimension. E.g., some evaluation goals do not necessarily need to involve academics in addition to practitioners, "it depends on the goal of the evaluation, you don't need academics if you only want to evaluate whether a units meets a certain criteria (eg. INSARAG [International Search and Rescue Advisory Group] classification, EU modeles classification)" (Rd. 2.).

Respondents also agreed upon the statement that evaluator teams shall be mixed in terms of ethnic and gender diversity, but the average was lower. Interestingly, this was controversial. One agreed, one made "no comment", and three declined. The common argument was that recruitment of evaluators must be oriented towards competence, not towards quotas. E.g., *"saving lives is more important than meeting certain quotas, therefore recruitment shall be oriented towards expertise and knowledge in related fields, and not gender or ethnicity." (Rd. 2.).* Elsewhere, an expert explicitly recommended to deploy "*a mix of national (location of exercise) and international evaluators" (Rd. 2.).*

Interestingly, agreement on the need to deploy evaluator teams per inject site was quite low, despite it may increase objectivity of assessments. But on the other hand, multiple evaluators per site increase both the resource input and imbalances in the ratio of trainees and evaluators. There was full agreement, with some elaborating on the statement: "*Sometimes there are 5-6 sites running simultaneously, while other times only 1 or 2, therefore it would be a waste of resources to have 12 evaluators for being able to cover all the sites for that 1-2 hours duration of the exercise when it is needed, when at the other 34 hours 6-8 evaluators would have been enough" (Rd. 2).* One pointed to the fact that this is, obviously, more a "*practical difficulty*", while it would actually be recommendable

to deploy at least one but at best more evaluators per inject site, *"It is obvious that you need evaluators in all exercise sites (if possible), and more than one of them in each site. The more information you collect, the better!"* (Rd. 2). The underlying idea of "the more information the better" needs critical reflection.

Further, there was rather strong support for an international pool of evaluators, which is reinforced by the high interest of prescribing to the pool (18 out of 19, 1 MV), as well as a pre-defined skill set for evaluators. There was slightly stronger agreement that such a skill set shall serve as minimum requirements rather than a soft guideline.

Establishment of landmarks: Further, experts agreed upon the need for the elaboration of the rules of the engagement, standard operation procedures and standardized evaluation tools. This indicates that a shared understanding of what a good evaluation and proper behaviour of evaluators is about is needed within the civil protection community. There was full agreement (n=5).

However, standardized procedures should consider a variety of, e.g., exercise formats *("Apart from these, I agree with all the topics about standardized procedures and training of evaluators. This should be different for tabletop, functional and full-scale exercises." (Rd. 1))* and respondents were quite reluctant on a too rigid idea of standardization.

Standardization may reduce flexibility and adaptability, thus the opportunity to select proper tools and methods for a given exercise – "*Evaluation tools are good to have and useful, but I don't think they need to be standardized, as this may compromised flexibility according to the needs of each exercise" (Rd. 1).* – and to maintain agency in complex and dynamic situations: "*Keep flexibility to adapt to learning objectives/style of exercise. Too standardized and rigid removes agency/responsibility. Rather invest in capacities and skills so people can adapt/design themselves." (Rd. 1*). This was supported in round 2 as well, "*I think a mixed approach is more appropriate here. The use of standardised tools (like EEGs) set standards, however, evaluators should also operate as ethnographers, in order to be able to report information which is not necessarily structured." (Rd. 2*). Further, too rigid standardization may further lead to the establishment of epistemic communities that exclude heterodoxic perspectives, thus reduce potential for critical discussion and innovation. One respondent provides the example of INSARAG: "*A bad example to my opinion is INSARAG classification: it is too much rule driven, leading to a community with group thinking" (Rd. 1*). But flexibility is needed, "*We must allow evaluators to describe how they see their role in a specific FSX after we have provided relevant info. We may find that we need different skills and perspectives that we initially thought of" (Rd. 2).* Thus, rather than rigid standards, skills and competencies, shall be taken as landmarks that guide recruitment of evaluators, but not as kick-off-criteria.

Holistic approach. Finally, there is a strong need to embed the evaluation into the broader disaster management cycle. This is related to what was discussed under the heading of "contextuality". Thus, evaluators must not just be able to conduct high-quality evaluations, but also to provide the system with findings, recommendations and so forth. "*In most exercises there is no link between the FSX and planning activities or how they were elaborated, which in some cases created an isolated environment for the exercise regardless of the area and conditions of intervention. One must be able to see the broader picture of emergency management." (Rd. 2)*

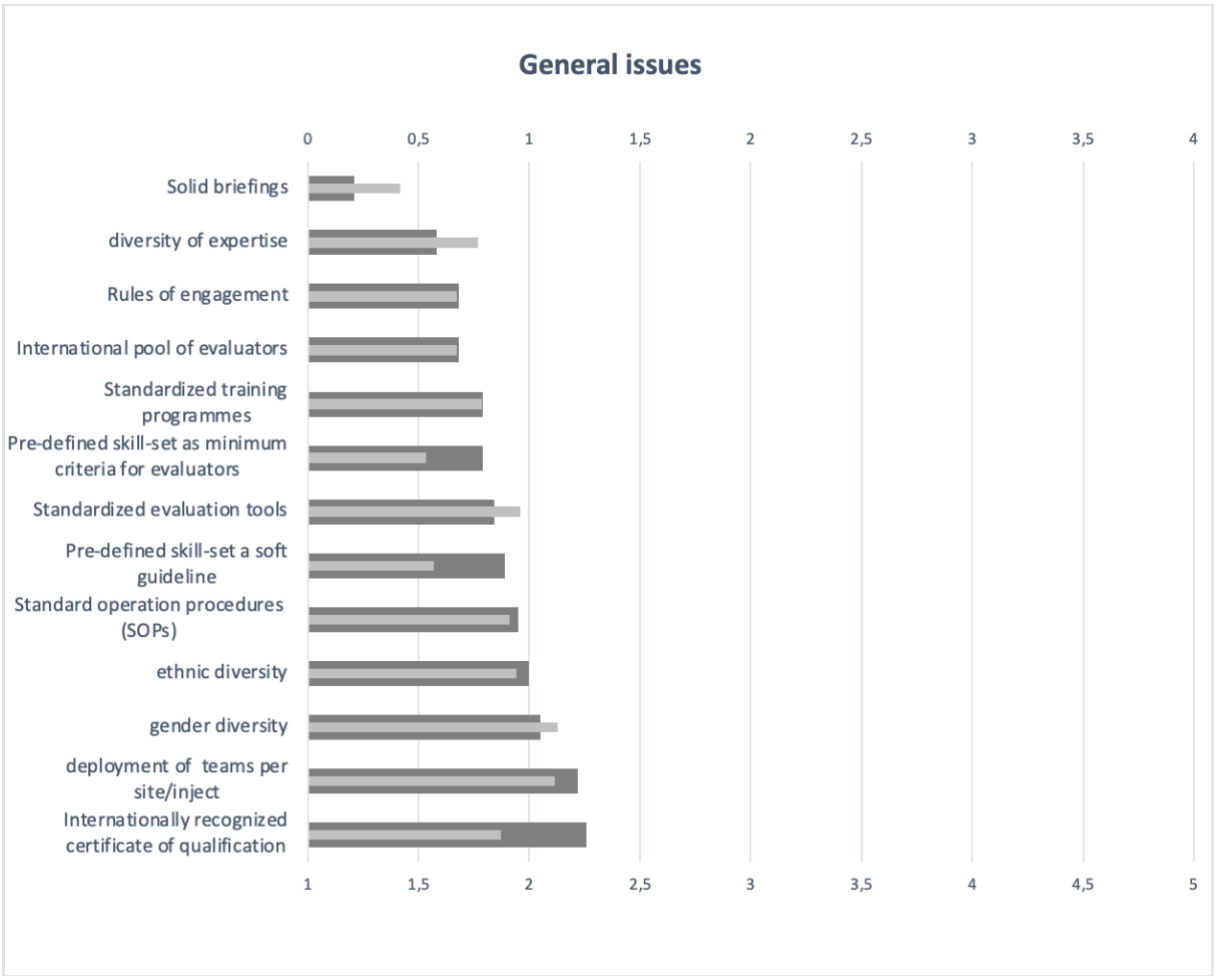

Figure 8: Agreement upon general issues regarding exercise evaluation

4 Conclusion

This deliverable provided detailed findings of the delphi study on the skill set of exercise evaluators. It was distinguished between skills and competencies, that evaluators *must* feature, *shall* feature or (depending on task) might feature. Generally speaking, being evaluator requires a well-defined mix of key competence, social and personal competence that may be considered "universal requirements" and a bundle of expert knowledge on subject matters, methods, etc., which are hardly to be defined in advance. For a detailed overview, see figure 9.

Due to a lack of a solid state of the art in the field of meta-evaluation of the evaluators, it was an exploratory analysis, and further research is needed to validate and complement the findings. For example, further research should strive for a comparison of the identified skill set and the skills and competencies of actual evaluators and their experiences. Further, it is required to elaborate on the diversity of evaluator profiles, e.g., to develop typologies and present their respective strengths and weaknesses. As mentioned in the previous sections, there is not "ONE evaluator", and it is more about the compilation of a heterogeneous team of evaluators.

The initial findings suggest that being evaluator may be considered as having a certain attitude and not just matching a pre-defined set of skills. This is important, as successful evaluation is considered as more than a well-elaborated evaluation. Rather, it depends on the situation dynamics in the evaluation process, interactions between evaluators and decision-makers, and the evaluator's personal traits (Sandermann 2011, 15f). However, the implementation of lessons-learned identified by the evaluators was beyond the scope of the project. Further research is needed on the conditions and broader context, in which evaluation contributes to a further development in the field.

Furthermore, it became clear that contextualisation of the required skills and competencies is important. Additional research is needed to elaborate on the link of certain exercises, formats, evaluation goals, concepts, etc. The study is limited in this regard, as it did not differentiate for each item along with criteria such as evaluation concepts, or even more detailed (e.g., certain tasks and roles) etc. for the reason of reducing the complexity of the questionnaire.

Further, it must be emphasized that the skill-set or competence profile for evaluators is not to be considered a fixed list. Adaptability and flexibility is important, especially regarding the specific requirements for any specific exercise.

Moreover, there is a strong need to elaborate on how to bring the skill-set into practice. Partly, this is implemented in the INEGMA-E² project. The findings on the skills and competences presented in this report will serve as base for the development of a training concept for evaluators and the conceptualization and establishment of the network of evaluators that may, in the future, support exercise planners and organizers in the recruitment and selection of evaluators. Beyond that, it is recommended to further elaborate on guidance for exercise planners and organizers on how to use the competence portfolio as a tool in the selection and recruitment process.

	MUST	SHOULD	MIGHT	Controversial
Basic requirements/ practical skills	Availability during exercise Physical attendance Analytical Skills Rules of Engagement Text Comprehension Safety Measures Contextuality Availability during preparation --- Minimum standards: English level B2-C1 Cognitive maturity (like secondary degree - Bachelor's level)	Ability to assess Organisational skills Abbreviations Formal terminology Colloquial language intercultural competence --- Previous exercise participation Affiliation to first responder organization Training degree Evaluation records Staff management Orientation Analog navigation Navigation systems Media competency Phonetic alphabet	Driving Licence First Aid Skills Basic local language skills	Affiliation to first-responder organization Problem orientation Solution orientation Local culture Creativity
Professional Knowledge Areas	Understanding of certain areas is important, but relate to the specific exercise	UCPM Tactics/strategies Disaster management system Exercise types Organisational culture Organisational structure Mandates Scenarios Resources Preparedness and contingency plans International civil protection measures EU ModEx		Legal framework of host nation International legal framework

Methodological	Observation Forms and checklists Computers	AAR/Debriefs Interviews Smartphone VCT Technical literacy Document analysis Messenger services Methods of analysis Evaluations apps Surveys Data processing Cameras Quantitative methods Qualitative methods	Radio communication	Workshops Methodological Meta-Competence
Social Competence	Communicativeness Constructive feedback Transparency Cooperativeness Tolerance Appreciative attitude Adaptability	Open to criticism Assertiveness	Intraversion Extraversion Subordination Competitiveness reservedness	Consensus-oriented attitude Conflict ability Empathy Rhetorical skills Willingness to compromise Leadership
Personal Competence	Objectivity Reliability Integrity Flexibility Critical self-assessment Learning capacity Holistic perspective Accuracy Questioning attitude Self-discipline Motivation Pragmatism Resilience Proactivism decisiveness	Idealism Pedantic Rapidity Self-determined working style Self-confidence	Impulsiveness externally determined working style prudency	Passivity

Figure 9: Overview of skills and competencies

5 References

Beerens, R.; H. Theler; B. Pelzer (2020): How can we make disaster management evaluations more useful? An empirical study of Dutch exercise evaluations. In: IJDRS. Doi: https://doi.org/10.1007/s13753-020-00286-7

Beerens, R.; K. Haverhoek-Mieremet (2020b): What do practitioners expect from an evaluation reports? In: IJES, doi: 10.1108/IJES-12-2019-0063

Beerens, R.; Theler, H. (2016): Scoping the field of disaster exercise evaluation. A literature overview and analysis. In: IJDRR, doi: 10.1016/j.ijdrr.2016.09.001

Bruns, Hendrik; Celikler, Jana; Jonitz, Hendrik (2022): A Framework for Evaluation Concept. INEGMA-E² Deliverable D2.2. Online available: https://civil-protection-knowledge-network.europa.eu/projects/inegma-e2 [2023-01-24].

Ruoslahti, Harri; Lonka, Harriet (2022). Evaluation methods. INEGMA-E² Deliverable D2.3. Online available: https://civil-protection-knowledge-network.europa.eu/projects/inegma-e2 [2023-01-24].

Karapidis, A.; B. Dworschak (2016): Study on Competence Management within the Union Civil Protection Mechanism Training Programme. Stuttgart: Fraunhofer IAO.

Sandermann, M. (2011): Die Bedeutung von Soft Skills für Evaluationsnutzungen. Wiesbaden: VS.

Swedish Civil Contingencies Agency. 2011. Handbook evaluation of exercises. Stockholm: Swedish Civil Contingencies Agency. https://www.msb.se/siteassets/dokument/publikationer/english-publications/evaluation-of-exercises.pdf

Attachment to D4.1

Delphi Expert Survey Questionnaire*

*) Limesurvey export

INEGMA-E2

International Network of Evaluators &
Guideline for a Methodological Approach in Exercise Evaluation

For further information see: https://civil-protection-knowledge-network.europa.eu/projects/inegma-e2

Dear Sir or Madam,

You have been identified as an expert in the field of exercise evaluation. Accordingly, we wish to invite you to participate in the delphi expert panel as a part of the project INEGMA-E^2 (UCPM) for the development of a skill-set for exercise evaluators.

The survey is part of the broader UCPM-funded project "INEGMA-E^2 - International Network of Evaluators & Guideline for a Methodological Approach in Exercise Evaluation."

The goal of the delphi panel is to build an informed expert consensus on the skills and competencies required for evaluation experts. It sets the base for improved recruiting and quality assurance mechanisms.

The survey is a multiple round panel. We will conduct two panels.
In the first round, you will be asked to make an initial assessment of the priority of various skills and further comments.
In the second round (December), you will be invited in a structured feedback, i.e. provided with the preliminary results and asked for further comments.

The questionnaire takes about 20 to 25 minutes. You can save your answers and continue later on ("resume later", in the top right corner). It consists of the following thematic blocks:

1) basic requirements (hard facts)
2) professional competence (areas of expertis)
3) methodological competence (skills and know-how)
4) social competence (skills for social interaction)
5) personal competence (skills for personal development)
6) metalevel questions (further issues regarding the framework for evaluation)
7) personal information

We would highly appreciate your willingness to share your expertise with us!

In case of any questions or concerns, please don't hesitate to contact us under:
office@inegma-e2.eu (project-related questions) or
sandra.pfister@dcna.at (questions related to the delphi panel)

Thank you for your valuable contribution and with best regards
the INEGMA-E^2 project consortium

There are 27 questions in this survey.

Basic Requirements

The following items refer to some potential basic requirements for evaluators.

Please indicate the level of priority of the following basic requirements for exercise evaluators. *

Please choose the appropriate response for each item:

	1 = high priority	2 = moderate priority	3 = low priority	4 = not required	I don't know/no opinion
availability at the time of the exercise.	○	○	○	○	○
availability at the time of exercise preparation.	○	○	○	○	○
ability to physically attend the exercise (e.g., international travelling)	○	○	○	○	○
affiliation to a practical organizations (e.g., first responder organization)	○	○	○	○	○
degree in any evaluation training programme	○	○	○	○	○
evaluation records from previous exercise evaluations.	○	○	○	○	○
confirmation of attendance of former civil protection exercises.	○	○	○	○	○
first aid qualification.	○	○	○	○	○
driving licence	○	○	○	○	○

Please indicate the minimum education status required for evaluators. *

Choose one of the following answers
Please choose **only one** of the following:

- ◯ no minimum education status required
- ◯ primary education (elementary school)
- ◯ lower secondary education (e.g., middle school, junior high)
- ◯ upper secondary education (e.g., high school, vocational education)
- ◯ university degree (Bachelor's, Master's, or equivalent)
- ◯ doctoral degree
- ◯ I don't know/no opinion

Make a comment on your choice here:

Please indicate the minimum required level(s) of language skills according to the CEFR (Common European Framework of Reference for Languages) for exercise evaluators? *

Please choose the appropriate response for each item:

	C2 Proficient User	C1 Proficient User	B2 Independent User	B1 Independent User	A2 Basic User	A1 Basic User	not required	I don't know/no opinion
working language of the exercise (i.e., English for EU-level exercises), general	◯	◯	◯	◯	◯	◯	◯	◯
working language of the exercise (i.e., English for EU-level exercises), reading skills	◯	◯	◯	◯	◯	◯	◯	◯
working language of the exercise (i.e., English for EU-level exercises), writing skills	◯	◯	◯	◯	◯	◯	◯	◯
working language of the exercise (i.e., English for EU-level exercises), speaking skills	◯	◯	◯	◯	◯	◯	◯	◯
working language of the exercise (i.e., English for EU-level exercises), listening skills	◯	◯	◯	◯	◯	◯	◯	◯
local language of the host nation, general	◯	◯	◯	◯	◯	◯	◯	◯
local language, reading skills	◯	◯	◯	◯	◯	◯	◯	◯
local language, writing skills	◯	◯	◯	◯	◯	◯	◯	◯
local language, speaking skills	◯	◯	◯	◯	◯	◯	◯	◯
local language, listening skills	◯	◯	◯	◯	◯	◯	◯	◯

Language proficiency levels according to the CEFR (Common European Framework of References for Languages)

A1 = Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help.

A2 = Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g. very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.

B1 = Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst travelling in an area where the language is spoken. Can produce simple connected text on topics which are familiar or of personal interest. Can describe experiences and events, dreams, hopes & ambitions and briefly give reasons and explanations for opinions and plans

B2 = Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.

C1 = Can understand a wide range of demanding, longer texts, and recognise implicit meaning. Can express him/herself fluently and spontaneously without much obvious searching for expressions. Can use language flexibly and effectively for social, academic and professional purposes. Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organisational patterns, connectors and cohesive devices.

C2 = Can understand with ease virtually everything heard or read. Can summarise information from different spoken and written sources, reconstructing arguments and accounts in a coherent presentation. Can express him/herself spontaneously, very fluently and precisely, differentiating finer shades of meaning even in more complex situations.

Would you like to add any comments on the section "basic requirements"?

E.g., would you suggest any variation between types of exercises (e.g., tabletop exercise, functional exercise, full-scale exercise) or the roles and tasks of evaluators?

Please write your answer here:

Professional Competencies

The following items refer to professional competence, which includes organization-, process-, task- and workplace-specific skills and knowledge in the context of disaster management, civil protection and exercise evaluation.

Please indicate the level of priority of the following areas of expertise required for exercise evaluators. *

Please choose the appropriate response for each item:

	1 = high priority	2 = moderate priority	3 = low priority	4 = not required	I don't know/no opinion
previous internatonal missions in civil protection	○	○	○	○	○
specific scenarios being trained (e.g., forest fires, CBRN, USAR etc.)	○	○	○	○	○
the disaster management system being evaluated (e.g., set of institutions and organizations involved in managing industrial accidents)	○	○	○	○	○
mandates (domain/roles) of the involved organizations and actors	○	○	○	○	○
formal structures (hierarchy etc.) of the involved organizations	○	○	○	○	○
tactics and strategies being trained in the exercise	○	○	○	○	○
resources (gear, equipment, etc.) being applied in the trained scenario	○	○	○	○	○
the organizational culture of the involved organizations ("informal structures", i.e. norms and values being shared)	○	○	○	○	○
preparedness and contingency plans within the system being trained	○	○	○	○	○
the international law relevant for civil protection exercises and missions	○	○	○	○	○
the European Civil Protection Mechanism	○	○	○	○	○

the legal framework of the host nation relevant for civil protection exercises and missions	○	○	○	○	○
different types of exercises (Tabletop, functional, full-scale, etc.)	○	○	○	○	○
EU Modules Exercises (ModEx)	○	○	○	○	○

Would you like to add any comments on the section "professional competence"?
E.g., would you suggest any variation between types of exercises (e.g., tabletop exercise, functional exercise, full-scale exercise) or the roles and tasks of evaluators?

Please write your answer here:

Methodological Competencies

The following items refer to methodological competencies, which include skills and know-how on tools and mehods, which can be applied flexibly and in different contexts to perform tasks.

Practical skills and competencies

Please indicate the level of priority of the following practical skills and competencies required for exercise evaluators. *

Please choose the appropriate response for each item:

	1 = high priority	2 = moderate priority	3 = low priority	4 = not required	I don't know/no opinion
analog navigation techniques (readings maps, etc.)	○	○	○	○	○
navigation systems and procedures (GPS, tracking, etc.)	○	○	○	○	○
Spatial awareness / orientation	○	○	○	○	○
local culture of the host community	○	○	○	○	○
intercultural competency	○	○	○	○	○
colloquial language (e.g., slang words) used in disaster management	○	○	○	○	○
formal terminology used in disaster management	○	○	○	○	○
common abbreviations used in disaster management	○	○	○	○	○
text comprehension (e.g., ability to read and understand contingency plans)	○	○	○	○	○
mastery of the international phonetic alphabet (alpha, bravo, charlie…)	○	○	○	○	○
staff management	○	○	○	○	○
organization skills	○	○	○	○	○

on-site security and safety regulations	○	○	○	○	○
rules of engagement (proper behavior and understanding of tasks, evaluators role)	○	○	○	○	○
media competency	○	○	○	○	○
analytical skills	○	○	○	○	○
contextuality (ability to understand interdependencies and complex relationships)	○	○	○	○	○
problem orientation	○	○	○	○	○
solution orientation	○	○	○	○	○
creativity/improvisation	○	○	○	○	○
ability to assess	○	○	○	○	○

Would you like to add any comments on the section "practical skills"?
E.g., would you suggest any variation between types of exercises (e.g., tabletop exercise, functional exercise, full-scale exercise) or roles and tasks of the evaluators?

Please write your answer here:

Methods expertise

Please indicate the level of priority of the following methods competencies required for exercise evaluators.
*

Please choose the appropriate response for each item:

	1 = high priority	2 = moderate priority	3 = low priority	4 = not required	I don't know/no opinion
quantitative research methods	○	○	○	○	○
qualitative research methods	○	○	○	○	○
interviews (codes of conducts, use of interview guidelines, etc.)	○	○	○	○	○
observations	○	○	○	○	○
surveys	○	○	○	○	○
document analysis	○	○	○	○	○
data processing	○	○	○	○	○
analytical methods (content analysis, statistics, etc.)	○	○	○	○	○

workshops	○	○	○	○	○
after action reviews (AAR)/debriefs/hot wash	○	○	○	○	○
methods meta-competency (ability to chose proper methods)	○	○	○	○	○

Please indicate the level of priority of the capacity to properly apply the following tools required for exercise evaluators. *

Please choose the appropriate response for each item:

	1 = high priority	2 = moderate priority	3 = low priority	4 = not required	I don't know/no opinion
technical and digital literacy	○	○	○	○	○
radio/satellite phone communication	○	○	○	○	○
messenger services (e.g., Signal, Telegram, etc.)	○	○	○	○	○
video conference tools	○	○	○	○	○
computers	○	○	○	○	○
smartphones/tablets	○	○	○	○	○
application and software supporting evaluation tasks	○	○	○	○	○
cameras/video recorders	○	○	○	○	○
forms and checklists	○	○	○	○	○

Would you like to add any comments on the section "methods expertise"?
E.g., would you suggest any variation between types of exercises (e.g., tabletop exercise, functional exercise, full-scale exercise) or roles and tasks of the evaluators?

Please write your answer here:

Social Competence

The following items refer to social competencies, which include skills and competencies that enable individuals to socially interact in a communicative, cooperative and self-organized way.

Please indicate the level of priority of the following social competencies and characteristics required for exercise evaluators. *

Please choose the appropriate response for each item:

	1 = high priority	2 = moderate priority	3 = low priority	4 = not required	I don't know/no opinion
leadership competency	○	○	○	○	○
assertiveness	○	○	○	○	○
subordination	○	○	○	○	○
adaptability	○	○	○	○	○
cooperativeness	○	○	○	○	○
competitiveness	○	○	○	○	○
conflict ability	○	○	○	○	○
willingness to compromise	○	○	○	○	○
consensus-oriented attittude	○	○	○	○	○
open to criticism	○	○	○	○	○
ability to give constructive criticism	○	○	○	○	○
appreciative attitude	○	○	○	○	○
tolerance and open-mindedness	○	○	○	○	○
empathy	○	○	○	○	○
transparency (e.g., regarding their own roles and actions)	○	○	○	○	○
communicativeness	○	○	○	○	○
rhetorical skills	○	○	○	○	○
reservedness	○	○	○	○	○
extroversion	○	○	○	○	○
introversion	○	○	○	○	○

Would you like to add any comments on the section "social competence"?

E.g., would you suggest any variation between types of exercises (e.g., tabletop exercise, functional exercise, full-scale exercise) or the roles and tasks of evaluators?

Please write your answer here:

Personal Competencies

The following items refer to personal competencies, which include the ability to self assessment and development.

Please indicate the level of priority of the following personal competencies and characteristics are required for exercise evaluators. *

Please choose the appropriate response for each item:

	1 = high priority	2 = moderate priority	3 = low priority	4 = not required	I don't know/no opinion
integrity/trustworthiness	○	○	○	○	○
reliability/responsible	○	○	○	○	○
resilience (capacity to absorb and deal with pyschologial and physiological burden)	○	○	○	○	○
decisiveness (decision-making capacity)	○	○	○	○	○
flexibility (ability to meet situational demands)	○	○	○	○	○
objectivity (ability to maintain a critical distance/to draw conclusions regardless of the personal stance)	○	○	○	○	○
questioning/critical attitude	○	○	○	○	○
proactivity	○	○	○	○	○
passivity/observant attitude	○	○	○	○	○
motivation/eagerness	○	○	○	○	○
learning capacity	○	○	○	○	○
critical self-assessment	○	○	○	○	○
impulsiveness	○	○	○	○	○
prudency	○	○	○	○	○
self-confidence	○	○	○	○	○
self-discipline	○	○	○	○	○
self-determined working style	○	○	○	○	○
externally determined working style	○	○	○	○	○
holistic attitude (considerative of "the big picture")	○	○	○	○	○
pedantic attitude (considerative of details)	○	○	○	○	○
rapidity (ability to perform tasks quickly)	○	○	○	○	○
accuracy (ability to perform tasks with diligence)	○	○	○	○	○
pragmatic attitude	○	○	○	○	○
idealistic attitude	○	○	○	○	○

Would you like to add any comments on the section "personal competencies"?
E.g., would you suggest any variation between types of exercises (e.g., tabletop exercise, functional exercise, full-scale exercise) or the roles and tasks of evaluators?

Please write your answer here:

Metalevel Questions

The following items refer to more general requirements for exercise evaluations.

For exercise evaluation in general, there is a need for… *

Please choose the appropriate response for each item:

	1 = strongly agree	2 = agree	3 = neither ... nor	4 = disagree	5 = strongly disagree	I don't know/no opinion
...an international pool of evaluators qualified for exercise evaluation	○	○	○	○	○	○
…a pre-defined skill-set as minimum criteria for evaluators	○	○	○	○	○	○
…a pre-defined skill-set a soft guideline for the selection of evaluators	○	○	○	○	○	○
…solid briefings prior to the exercise to be evaluated.	○	○	○	○	○	○
…standardized training programmes for exercise evaluators.	○	○	○	○	○	○
…internationally recognized certificate of qualification for exercise evaluators.	○	○	○	○	○	○
…standard operation procedures (SOPs) in exercise evaluation.	○	○	○	○	○	○
...rules of engagement for exercise evaluators.	○	○	○	○	○	○
...standardized evaluation tools.	○	○	○	○	○	○
…the deployment of evaluator teams per site/inject (rather than individual evaluators)	○	○	○	○	○	○
…diversity of involved evaluators in terms of skills and expertise	○	○	○	○	○	○
…diversity of involved evaluators in terms of gender	○	○	○	○	○	○
…diversity of involved evaluators in terms of nationality	○	○	○	○	○	○

Would you like to add any comments on the section "metalevel questions"?
E.g., would you suggest any variation between types of exercises (e.g., tabletop exercise, functional

exercise, full-scale exercise) or the evaluation concept and goals?

Please write your answer here:

Personal information

Please indicate your role(s) related to civil protection exercises. *

ℹ Check all that apply
Please choose **all** that apply:

- ☐ I'm a researcher in the area of civil protection exercises (please note your research topics in the box below).
- ☐ I'm an evaluator of civil protection exercises (please note the type(s) of exercises and the tasks and goals of the evaluations in the box below).
- ☐ I'm an organizer of and/or trainer within civil protection exercises (please note the type(s) of exercises and your particular tasks in the box below).
- ☐ I have participated in civil protection exercises (please note the type(s) of exercises and your particular tasks in the box below).
- ☐ Else (please note further details in the box below).

Further details on your role(s) related to civil protection exercises:

Please write your answer here:

Please select your gender. *

ℹ Choose one of the following answers
Please choose **only one** of the following:

- ◯ male
- ◯ female
- ◯ diverse

What is your age? *

ℹ Only numbers may be entered in this field.
Please write your answer here:

Are you affiliated to an emergency response organization (e.g., fire brigade, paramedical services, emergency operation centers, etc.)?
(If yes: Please note in the comments, which kind of organization) *

ℹ Choose one of the following answers
Please choose **only one** of the following:

- ◯ Yes
- ◯ No

Make a comment on your choice here:

What is your country of residence? *

Please write your answer here:

Please indicate the level of your English proficiency according to the CEFR (self-assessment).

ℹ Choose one of the following answers
Please choose **only one** of the following:

- ◯ A1 - Basic User
- ◯ A2 - Basic User
- ◯ B1 - Independent User
- ◯ B2 - Independent User
- ◯ C1 - Proficient User
- ◯ C2 - Proficient User

Language proficiency levels according to the CEFR (Common European Framework of References for Languages)

A1 = Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help.

A2 = Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g. very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.

B1 = Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst travelling in an area where the language is spoken. Can produce simple connected text on topics which are familiar or of personal interest. Can describe experiences and events, dreams, hopes & ambitions and briefly give reasons and explanations for opinions and plans

B2 = Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.

C1 = Can understand a wide range of demanding, longer texts, and recognise implicit meaning. Can express him/herself fluently and spontaneously without much obvious searching for expressions. Can use language flexibly and effectively for social, academic and professional purposes. Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organisational patterns, connectors and cohesive devices.

C2 = Can understand with ease virtually everything heard or read. Can summarise information from different spoken and written sources, reconstructing arguments and accounts in a coherent presentation. Can express him/herself spontaneously, very fluently and precisely, differentiating finer shades of meaning even in more complex situations.

One of the further project goals of INEGMA-E2 is to develop a network of evluators. The network will serve as an expert pool for the future selection and invitation of exercise evaluators.

I want to pre-register for the network of evaluators.
(If yes: Please provide your email-address in the comments). *

ℹ Choose one of the following answers
Please choose **only one** of the following:

- ◯ Yes
- ◯ No

Make a comment on your choice here:

**) Please note that if you agree with being registered, we will store your provided contact details for that purpose in an extra(!) file (not linked to your response data).*

Thank you for your participation in our delphi expert panel as a part of the project INEGMA-E^2 (UCPM). We highly appreciate your valuable contribution!

After having evaluated the results from round #1, we will invite you to participate in round #2: structured feedback (expected: December).

In case of any questions, please don't hesitate to contact us under:
office@inegma-e2.eu (project-related questions) or
sandra.pfister@dcna.at (questions related to the delphi panel)

With best regards
the INEGMA-E^2 project consortium

20.12.2022 – 10:49

Submit your survey.
Thank you for completing this survey.