A Metrology based Approach for measuring the Social Dimension of Cognitive Trust in Collaborative Networks

Andrade-Garda, Javier; Anguera, Áurea; Ares-Casal, Juan; García-Vázquez, Rafael;

Lara, Juan-Alfonso; Lizcano, David; Rodríguez-Yáñez, Santiago; Suárez-Garaboa,

Sonia

This paper addresses the measurement of the social dimension of cognitive trust in collaborative networks. Trust indicators are typically measured and combined in literature in order to calculate partners' trustworthiness. When expressing the result of a measurement, some quantitative indication of the quality of the result—the uncertainty of measurement—should be given. However, currently this is not taken into account for the measurement of the social dimension of cognitive trust in collaborative networks. In view of this, an innovative metrology based approach for the measurement of social cognitive trust indicators in collaborative networks is presented. Thus, a measurement result is always accompanied by its uncertainty of measurement, as well as by information traditionally used to properly interpret the results: the sample size, and the standard deviation of the sample.

Keywords: collaborative network; measurement; metrology; social cognitive trust indicator; uncertainty of measurement

1 Introduction

Nowadays it is commonly accepted that, in a rapidly globalizing world, enterprises' structures and processes should evolve in order to deal with the globalization of the economy, the rapid growth of information technologies and the increase of competitiveness (García et al. 2016). In this context, the alliance of organizations is an essential mean for attending customers and business opportunities. This way, enterprise cooperation can be defined as an agreement between two, or more, independent enterprises that, not merging but joining or sharing some of their capabilities and/or resources, can establish some kind of interrelation to increase their competitive advantages (Andrade et al. 2015). Thanks to the advance of the information and communications technologies (ICTs), such cooperation could take the form of a collaborative network (CN).

Following Camarinha-Matos and Afsarmanesh (2005), a CN is constituted by a variety of entities (e.g., organizations and people) that are largely autonomous, geographically distributed, and heterogeneous in terms of their: operating environment, culture, social capital, and goals. Nevertheless, these entities collaborate to better achieve common or compatible goals, and whose interactions are supported by computer network. In CNs, collaboration is an intentional property that derives from the shared belief that together the network members can achieve goals that would not be possible or would have a higher cost if attempted by them individually. Thus, cooperative work creates several advantages such as complementing individual abilities, skills, and knowledge;

and can help manage difficult and complex tasks that cannot be easily addressed by a single participant (Nonose et al. 2016, 2014). The advantages provided by CNs are appreciated by many types of organization including industry, business and public sector organisations such as government, education and health (Camarinha-Matos 2014). As a consequence, a large variety of CNs have emerged including virtual organizations, virtual enterprises, dynamic supply chains, professional virtual communities, collaborative virtual laboratories, etc. (Camarinha-Matos and Afsarmanesh 2005).

Once they appeared, CNs gained much attention from many academics, as well as the trust concept that is involved in their entire life cycle (Camarinha-Matos and Afsarmanesh 1999). In this regard, as stated in (Dunn 2000), the literature has identified two aspects to trust: a cognitive element in which trust is the result of a rational calculation by the trustor about how the trustee will behave in the future, and an emotional element in which trust is the product of a strong positive affection between the two individuals. Most business relations are based on cognitive trust, which is obtained as the result of the cognitive process of building trust between partners, whereas emotional trust is the basis for intense personal relationships, such as love and friendship (Dunn 2000). Therefore, this paper focuses on the cognitive trust (hereinafter referred to as trust), which is the important one for establishing effective business relationships in CNs, since it promotes cooperation and coordination (Mayer and Gavin 2005). This is because, in general, team members are much more willing to cooperate with and coordinate with individuals they believe are competent (Dirks 1999). In fact, Fang et al. (2014) identified the level of trust among members as a key factor that influences the effectiveness of the knowledge processes in virtual teams. Trust is considered as both an important pre-condition as well as a result of collaboration (Rusman et al. 2013). In online environments it is based on beliefs in the trustworthiness of a trustee (Gefen et al. 2008). According to Gambetta (1988), trust is a particular level of the subjective probability with which an agent will perform a particular action, both before [we] can monitor such action (or independently of his capacity to ever to be able to monitor it) and in a context in which it affects [our] own action. As Chen et al. (2015) indicate, trust is a catalyst that facilitates strategic business interactions and knowledge sharing among independent firms. The discussion on significance of trust for CNs started with the work of Charles Handy (Yasir et al. 2014). Handy (1995) observed that establishing trust among members is pivotal for the success of a virtual organization. Other researches (e.g., Jarvenpaa and Leidner 1999; Panteli and Duncan 2004; Kanawattanachai and Yoo 2007) addressed the importance of trust in global teams linked by ICTs. Specifically, Jarvenpaa and Leidner (1999) argue that trust is maximally important in new and temporary organizations, because it acts as a substitute for the traditional mechanisms of control and coordination. Trust acts as a buffer that facilitates the agreement and execution of transactions (Kasper-Fuehrer and Ashkanasy 2001). Therefore, it is now an established fact that trust among parties is extremely important for a successful virtual relationship (Yasir et al. 2014). The term trust indicator/factor is used to refer to the interest issues that contribute to trust among parties.

Examples of such trust factors are numerous. For instance, as shown in (Barki et al. 2015), Mayer et al. (1995) identified ability, benevolence, and integrity to be three key characteristics of trustworthiness that help explain much of the within-truster variation observed in trust. As shown in (Fan et al. 2011), this is also supported by, among others, Jarvenpaa et al. (1998), and Greenberg et al. (2007). Similarly, McKnight et al. (2002) propose trust consisting of the beliefs of integrity, benevolence, and competence that the trustor has in the trustee. Xu et al. (2016) analyse the role of the three trust beliefs (McKnight model) in predicting two distinct outcomes: satisfaction and purchase

behaviour. Nakayama et al. (2006) pointed out that trust is related to competence, loyalty and receptiveness; Fan et al. (2009) emphasize the importance of collaboration satisfaction. Many other examples of trust factors can be found in literature (e.g., Sabater and Sierra 2002; Lavrac et al. 2007; Haller 2008; Msanjila and Afsarmanesh 2009; Leina and Tiejun 2012; Tian and Wang 2012): honour, operational costs, auditing frequency, innovation, risk willingness, punctuality, partnership, etc. In fact, Haller's taxonomy encompasses 146 trust indicators of varying kinds (Haller 2008).

To somehow summarize trust indicators, Paul and McDaniel (2004) present three forms of trust related to virtual organizations: calculative, competence and relational trust. Other proposals on forms of trust can be found in, for example, (Kanawattanachai and Yoo 2007), (Lambrechts et al. 2009), (Msanjila and Afsarmanesh 2010), and (Berry 2011). In a later study, Hardwick et al. (2013) found that it is possible to readily distinguish between the dimensions of trust based upon technical capability and trust built from more personal dimensions. This is not a new idea, since Sako considered in 1992 two basic forms of trust: competence (technical based) and goodwill (social based) to describe the main characteristics of most trust types (Sako 1992). Note that both forms of trust refer to cognitive trust as they involve the measurement of trust factors in order to compute the trustee's trustworthiness. Trust can be very dynamic, increasing or decreasing based on whether an individual fulfils or fails to fulfil commitments (Robert et al. 2009), so it is necessary to adequately measure it over time.

Although both technical and social based trust can be difficult to measure, technical based trust can be obtained many times from raw or statistical project data (e.g., delivery time, costs, and product defects) as well as via other objective methods. However, social based trust depends on opinion and is obtained through interviews and/or questionnaires, where subjectivity is a natural and key factor to be necessarily considered. Thus, although special measurement capabilities are necessary, in complex work environments such as CNs, the social aspects of performance—i.e., knowing about the relationships, collaborations and communications, of the participants with each other—are as important as any other aspect (Farrington-Darby and Wilson 2009).

Both social and technical trust indicators, as well as their measurement, are considered in the literature at different levels. The basic trust indicators (usually provided by the decision maker or calculated on the basis of historical data) are combined in (complex) formulas in order to obtain a final value of the trustworthiness of each partner, even going through several intermediate layers of indicators (calculations).

For example, in (Lavrac et al. 2007) trust in CNs is modelled using a hierarchical tree of decision criteria with two branches *reputation* (REP) and *collaboration* (COL). Thus, the trust of agent X in agent Y is calculated by a weighted sum of REP(Y) and COL(X,Y), where REP(Y) is the average of values of the basic input attributes (activity, punctuality, reliability, partnership, risk willingness, and economical situation) that are gathered via a questionnaire-based approach or via social network analysis, and COL(X,Y) is a number between 0 and 3 representing the frequency with which X and Y work together. In the proposal in (Tian and Wang 2012) *quality* (QL) and *capacity* (CP) trust indicators are considered and finally combined in order to obtain the global reputation of a partner ($C = \alpha QL + \beta CP$, $\alpha + \beta = 1$). The values of the trust indicators in both categories are provided by the decision maker knowing that CP indicators mainly depend on subjective evaluation and QL indicators are primarily based on past experience. In the model proposed in (Msanjila and Afsarmanesh 2009) five trust perspectives (Technological, Social, Structural, Managerial, and Economical) are proposed. Simple and composite (i.e., calculated as a function of others) trust factors are combined in order to compute the score of each trust perspective. Simple trust factors are

known factors provided by the decision maker. In the ReGReT system trust considers both individual and social reputation (Sabater and Sierra 2001). The individual reputation focuses on the direct interactions, in contrast to the social reputation, which also depends on opinions of third party agents. The information regarding direct interactions is always context-dependent and is linked to a certain behavioural aspect. Individual reputation at a certain time can be calculated by weighted mean of impressions (i.e., opinions from other partners) given more relevance to recent impressions. The same approach is adapted to calculate social reputation. Haller's approach indicates that a trust factor can be based on observations of a continuous or discrete variable x (Haller 2008). The distribution's expectation value is returned as the trust factor value. It also defines an observation time period, which delimits the maximal time window to look into the past. In this case, as in the ReGReT system, more relevance is given to more recent observations. As a last example, Fan et al. (2011) consider two dimensions of reputation and collaboration to estimate a team trust level. Each dimension is composed of several trust factors. The assessment information of these trust factors can be obtained either directly from team leaders (members) or indirectly from statistical data. Thus, a weighted assessment result of one member to another partner is obtained, as well as the performance of each member and the overall performance of the virtual team.

In summary, the main objective in all the above-mentioned proposals is to obtain the exact value (i.e., the precise and unique value) of each trust indicator at any level. It may be acceptable for technical trust indicators, where the result can be obtained via relatively objective methods. However, the subjectivity inherent in social trust indicators makes it necessary to explicitly take it into account and provide some quantitative indication of the uncertainty of the result, in order to allow users to assess its suitability. Note that a partner expressing her/his opinion about the same partner and regarding the same social trust factor at different times does not necessarily give the same result, even for the same collaboration. In this situation, some representation of the human/environmental influence on the values of social trust indicators is needed (i.e., the measurement should be accompanied by its uncertainty of measurement). Some attempts in this direction can be found in the literature. For example, the credibility factor of the feedback source is used to compute a weighted average of ratings (see e.g., Xiong and Liu 2003; Ruohomaa and Kutvonen 2010) in order to express some kind of uncertainty of the result. Similarly, the partner´s own trustworthiness is used as weighting factor in, for example, (Kamvar et al. 2003) meaning that the opinions of more reputable partners are more “valuable”. Those are concepts closely linked to recommendation systems and the detection of malicious behaviour such as positive and negative deceptive opinions (e.g., Simone et al. 2012; Hernández-Fusilier et al. 2015). Thus, these elements can be incorporated into the calculation process (e.g., as weighting items in order to obtain a weighted result) but their purpose is by no means to deal with the mathematical concept of uncertainty of measurement. A different approach to the problem can be found in the ReGReT system (Sabater and Sierra 2001) where in order to know the reliability of an individual/social reputation value the model considers two elements: the number of impressions used to calculate the reputation value, and the rating deviation of the impressions in a weighted sum. This approach is similar to that used in the Sporas system (Zacharia and Maes 2000). Similarly, the Haller’s approach also uses the variance to indicate the uncertainty inherent to the results (Haller 2008). However, none of them properly solve the problem addressed in this paper since they only provide the sample size and dispersion of individual observations (opinions), which are necessary values for the calculation of the uncertainty of measurement, but not properly the value of the uncertainty of measurement.

The word "uncertainty" means doubt. Thus, in its broadest sense, uncertainty of measurement means doubt about the validity of the result of a measurement (Bell 2008). In other words, the uncertainty of the result of a measurement reflects the lack of exact knowledge of the value of the measurands (JCGM 2008, 2012). The uncertainty of measurement is addressed by metrology. As stated in (JCGM 2012), metrology is the science of measurement and its application, and it includes all theoretical and practical aspects of measurement, whatever the measurement uncertainty and the field of application. In fact, the Joint Committee for Guides in Metrology (JCGM) publishes the Guide to the Expression of Uncertainty in Measurement (GUM) (JCGM 2008), which represents the most comprehensive and currently accepted reference for evaluating and expressing uncertainty in measurement. Following this guide, the uncertainty of measurement is part of the measurement result, which is usually represented as $R \pm U$, where R is the most likely outcome and U is the uncertainty of measurement associated with it (JCGM 2008, 2012). Thus, R is only an approximation or estimate of the value of the measurand and, therefore, only makes sense if it is accompanied by U. This approach is already used in other research areas (e.g., Chemistry, Physics, Sociology, and Psychology). The main purpose of this work is precisely to define a metrology based approach for measuring any social (cognitive) trust indicator in CNs.

This paper is organized as follows. Section 2 describes the data model behind the calculation method, which is presented in Section 3. An example in order to illustrate the method is presented in Section 4. Finally, Section 5 presents the most relevant conclusions and future work.

2 Data model

As it was said previously, social trust indicators depend on partners' opinion, which is formed through experience. In order to collect each partner's opinion on a particular trust indicator it is necessary to conduct interviews and/or questionnaires (let us refer to questionnaires henceforth, as nothing is different when data is collected through interviews).

Thus, the values of such trust indicators are calculated from historical collaboration data (questionnaire results), or they are directly provided by the decision maker, and combined somehow in order to obtain the values of higher level trust indicators, even obtaining an absolute and unique trust value for each partner. In this regard, what emerges from the analysis of the proposals presented in Section 1 is that trust management in CNs works with trust indicators at different levels. Specifically, there are two different but related perspectives on the trustworthiness of partners in CNs: (i) trust of one partner to another partner (*peer-to-peer*), and (ii) the general trustworthiness of a partner (usually based on the trust from others; i.e., based on *peer-to-peer* values). The data model presented in this section addresses both perspectives for each social trust indicator. As a consequence, a two-level data model, composed of the *peer-to-peer* and *global* levels, is proposed for social trust indicators. In this model, each partner calculates and maintains the *peer-to-peer* data about the value of the partners she/he/it has collaborated with in the considered trust factors (e.g., P_1 maintains the data about the *Integrity* of P_2 and P_3, as they have collaborated with P_1, and such data is based on such collaborations). This *peer-to-peer* data is combined by the "system" in order to obtain the *global* values (also maintained by the "system") for each partner in each trust factor (e.g., *global* data in the *Integrity* of a partner calculated from the *peer-to-peer* values of *Integrity* for that partner, maintained by each other partner with which she/he/it has collaborated).

This way, the *peer-to-peer* data level refers to (i) the results gathered from questionnaires on the trust indicators completed by the partners after each collaboration and, (ii) the combination of these results in order to obtain the value of each partner on each trust indicator from the point of view of the other partners.

Thus, for each CN it is necessary to firstly decide the trust factors that contribute to the trustworthiness of partners, and that, therefore, need to be measured. Once the list of trust factors has been decided, data is collected as follows: after each collaboration between two partners P_i and P_j (e.g., a joint task), P_i answers a questionnaire expressing her/his/its opinion about the score of P_j in each trust factor and regarding the collaboration that has just ended. P_j does the same for P_i. Thus, for example, if five trust factors are being considered then a collaboration between P_i and P_j generates ten questionnaire results. To minimize response errors, questionnaires are crafted in accordance with best practices. Thus, questions are worded in such a way that: they are simple, direct, and comprehensible; they are specific and concrete (rather than general and abstract); ambiguous words, double-barrelled questions, and negations are avoided; questions on the same topic are grouped together; etc. Further recommendations about optimal questionnaire design based on the common wisdom and on a review of the methodological research can be found in (Krosnick and Preser 2009). For responses, a rating scale has been designed. In this case, it ranges from 1 to 6, since this is the scale used in (Lavrac et al. 2007) to measure trust factors. However, it could also be possible to choose any other point scale, knowing that the length of scales can impact the process by which people map their attitudes onto the response alternatives. Further recommendations about rating scales can also be found in (Krosnick and Preser 2009). The responses to the questions in each questionnaire are combined in order to obtain the result value representing the opinion of the respondent on the issue in question (e.g., opinion of P_i about the integrity of P_j in their last collaboration). In order to simplify the interpretation of the results and to maintain consistency in the measurement, the questionnaire results in this paper (see Section 4 for illustrative example) also ranges (as the responses to the questions in the questionnaires) from 1 to 6. This way, a simple response average is used to obtain the questionnaire result. However, again, any other rating scale and calculation method could be used for questionnaire results. For example, in (Ashtiani et al. 2015), based on a verbal judgement of “very good”, “good”, “fair”, “poor”, and “very poor” provided by the respondent for each trust factor, the score of each trustee (i.e., partner on which the opinion is being expressed) in each trust factor is calculated by a fuzzy analytical hierarchical process. The approach is applied to service selection in the context of service-oriented environments, but it could be adapted to partner selection in the context of CNs. The important thing here is that a final numerical value for the opinion of each participant about each other on each trust factor is obtained. Whatever the questionnaires, the rating scale, and the process used to obtain the results, these results, just like any other measurement, must be accompanied by the uncertainty of measurement, the calculation of which is the main objective of this paper (Section 3). In other words, the aim of this paper is not to make a proposal on how to collect the partners’ opinions about each other but on how to work with the data obtained in order to provide a better representation of the meaning of the measurement.

This way, to put together the useful information items identified in literature and the new information requirements (i.e., the uncertainty of measurement), each item in a *peer-to-peer* data structure is defined as a tuple of the form (TF, C, (R, U, S, N)), where:

- TF is the considered social trust factor.
- C is the activity-related context. This activity-related context is usually considered in literature (e.g., Sabater and Sierra 2001; Hermoso et al. 2007; Afsarmanesh et

al. 2011), but it is optional (can be Ø). In this regard, it may be desirable to measure a trust factor from a general point of view (i.e., regardless of the type of activity performed by the partner) or to measure the trust factor subject to a particular working context (e.g., when the partner is performing testing activities). Respectively, *C* = Ø and *C* = "*Testing*". Therefore, the context is conditioned by the practical environment in which this approach to the measurement of trust factors in CNs is being applied. Thus, for example, in the information systems development domain, the environment could define, among others, the values of "*Analysis*", "*Design*", "*Coding*", and "*Testing*" for the context. That is to say, it is possible to measure a social trust factor (e.g., integrity) or a social trust factor in an activity-related context (e.g., integrity in testing activities).

- For each social trust factor (and context, if it applies), four values are provided:
  - *R* is the value of the social trust factor itself.
  - *U* is the uncertainty of measurement (its calculation procedure is presented in Section 3).
  - *S* is the standard deviation of the sample.
  - *N* is the number of collaborations on which the results are based.

These two-last mentioned values (*S* and *N*) are really useful because of the following. Think about a social trust factor with value 4 (i.e., *R* = 4). The "4" seems more stable if it is based on 20 observations and less stable if it is based on 5. Likewise, it seems more stable if it comes from the average of observations {4, 3, 4, 5} and less stable if it comes from the average of observations {2, 6, 2, 6}. Thus, these values help to increase the knowledge about cooperation which, as stated in (Nonose et al. 2016), contributes to the improvement of team performance (i.e., of CN performance).

Taking all the above into account, Fig. 1 shows an example of *peer-to-peer* data structure of a given partner P_1. In this case, the trust factors of *Integrity*, *Communication*, and *Assertiveness* are shown only for illustrative purposes, without entailing any assumption by default regarding their importance. In every particular case in which this measurement approach is being applied it will be necessary to define the specific trust factors to be considered, due both to the characteristics of the project and of the partners; the data structures here proposed are generic and, therefore, equally applicable to any trust factor.

$$P_2: \begin{bmatrix} (Integrity, \emptyset, (4.33, 0.56, 0.887, 12)) \\ (Communication, Testing, (\dots, \dots, \dots, \dots)) \\ (Assertiveness, \emptyset, (\dots, \dots, \dots, \dots)) \\ \dots \end{bmatrix}$$

$$P_3: \begin{bmatrix} (Integrity, \dots, (\dots, \dots, \dots, \dots)) \\ (Communication, \dots, (\dots, \dots, \dots, \dots)) \\ (Assertiveness, \dots, (\dots, \dots, \dots, \dots)) \\ \dots \end{bmatrix}$$

$$\dots$$

Fig 1. Example of *peer-to-peer* data structure

Thus, each partner P_i maintains a *peer-to-peer* data structure with *m* items, where *m* is the number of partners with which P_i has collaborated. Each of these items contains, among other data, P_i's opinion about the value of P_j on each social trust indicator. These opinions are calculated through the combination of the corresponding questionnaire results. This two-sided *peer-to-peer* data structure is used because opinion does not fulfil

the symmetric property. That is to say, P_i's opinion about P_j does not have to be the same as P_j's opinion about P_i even for the same collaboration and social trust indicator.

Going up one level in the proposed two-level data model, a partner could be interested to know, for example, the integrity of another partner, not just her/his own opinion or the opinion of another given partner on that issue. To make this possible, the *global* level values are maintained for each partner regarding each social trust factor. Following the idea expressed in, for example, (Kamvar et al. 2003) these values are derived from the corresponding *peer-to-peer* values. Obviously, note that, as in any case when a new variable is obtained by combining others, the uncertainty of measurement (*U*) has to reflect the uncertainty of measurement propagated from the corresponding *peer-to-peer* values. The same goes for the standard deviation (*S*). Fig. 2 shows an example of *global* level data structure.

$$P_1: \begin{bmatrix} (Integrity, \emptyset, (3.95, 0.21, 1.23, 7)) \\ (Communication, \dots, (\dots, \dots, \dots, \dots)) \\ (Assertiveness, \dots, (\dots, \dots, \dots, \dots)) \\ \dots \end{bmatrix}$$

P_2: [...]

P_3: [...]

...

Fig 2. Example of *global* data structure

This two-level data model constitutes more complete information than that submitted before and, for example, allows partners to better specify their collaboration requirements in CNs. Thus, for instance, in the creation stage, a partner could indicate that she/he is only interested in collaborating with partners with a global value of integrity of at least 3 in testing activities, a related uncertainty of measurement of at most 0.5, a standard deviation of the sample of at most 1, and that she/he only believes in results based on at least 10 *peer-to-peer* values. The same could be done with *peer-to-peer* values and restrictions on both levels can also be combined.

3 Metrology based computation procedure

According to GUM, uncertainty can be expressed in three different ways, depending on how it was obtained and the desired confidence interval. Thus, if *x* is the estimate of the variable *X*, the following applies:

(1) Standard uncertainty, *u(x)*: uncertainty of the result of a measurement expressed as a standard deviation.

(2) Combined standard uncertainty, $u_c(x)$: standard uncertainty of the result of a measurement when that result is obtained from the values of a number of other quantities.

(3) Expanded uncertainty, *U(x)*: quantity defining an interval about the result of a measurement that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand. The coverage factor (*k*) is the numerical factor used as a multiplier of the combined standard uncertainty in order to obtain an expanded uncertainty. Formally, $U(x) = k * u_c(x)$.

The method of evaluation of uncertainty depends on how the value of the measurand is estimated, and there are two types:

(1) Type A evaluation: method of evaluation of uncertainty by the statistical analysis of a series of observations.
(2) Type B evaluation: method of evaluation of uncertainty by means other than the statistical analysis of a series of observations.

The previous will enable the data structures specified in the two-level data model presented in Section 2 to be progressively fulfilled through a three steps metrology based process, which manages well known mathematical concepts (e.g., sample size, average, variance, and standard deviation).

3.1 Procedure for peer-to-peer level

The first thing to do is logically to define what is to be measured. In this case, measurands at *peer-to-peer* level (Section 2) can be defined as, for example, *X = "P_i's view about the TF (score) of P_j in C"*. This definition considers the conditions of measurement (required by GUM), that are expressed here in the form (P_i, P_j, *TF*, *C*) where P_i, P_j, and *TF* are respectively the involved partners and social trust factor, and *C* is the activity-related context (optional). That is to say, for example, variable *X_1 = "P_1's view about the Integrity of P_2"* can be defined under the conditions of measurement (P_1, P_2, Integrity, Ø), and variable *X_2* = "*P_1's view about the Integrity of P_2 in testing activities"* can be defined under the conditions of measurement (P_1, P_2, Integrity, Testing).

3.1.1 Step 1: Standard uncertainty calculation

As it was indicated in Section 1, the values of the basic (*peer-to-peer* level) social trust indicators are usually provided by the decision maker or calculated on the basis of historical data (i.e., a set of questionnaire results). The first case matches with Type B evaluation of uncertainty, and the second one matches with Type A evaluation. Although in most cases social trust indicators are based on available historical data (i.e., after each collaboration, each involved partner answers questionnaires about each other with whom she/he has collaborated), the two types of evaluation will be presented next, starting with type A evaluation.

Firstly, the estimate of each variable X_i must be defined. According to GUM, in most cases, the best available estimate of the expectation or expected value μ_q of a quantity *q* that varies randomly, and for which *n* independent observations q_k have been obtained under the same conditions of measurement, is the arithmetic mean or average of the *n* observations:

$$\bar{q} = \frac{1}{n}\sum_{k=1}^{n} q_k \tag{1}$$

This applies to each variable under Type A evaluation in this paper. That is to say, each variable that depends on *n* independent observations (i.e., questionnaires results) that have been obtained under the same conditions of measurement. Thus, each input estimate $x_i = \bar{X}_i$ and its associated standard uncertainty $u(x_i)$ are obtained from a frequency based distribution of possible values of the input quantity X_i (i.e., based on a series of observations $X_{i,k}$ of X_i).

In this regard, when the estimate of the measurand is the average of a series of observations, the focus of interest is the quantification of how well the mean estimates

the expectation μ_q of q, which may be used as a measure of the uncertainty of the mean (i.e., the uncertainty of the measurement). The values to be obtained are the experimental variance of the mean and the experimental standard deviation of the mean.

To this end, firstly the sample variance of the observations, which estimates the variance of the probability distribution of q [given by Eq. (2)], and its positive square root (i.e., the sample standard deviation) must be calculated. These values characterize the variability of the observed values q_k and they make it possible to calculate the experimental variance of the mean and the experimental standard deviation of the mean, equal to its positive square root [see Eq. (3)].

$$s^2(q) = \frac{1}{n-1}\sum_{k=1}^{n}(q_k - \bar{q})^2 \tag{2}$$

$$s^2(\bar{q}) = s^2(q)/n, \;\; s(\bar{q}) = s(q)/\sqrt{n} \tag{3}$$

Thus, for an input quantity X_i (e.g., *"P_1's view about the Integrity of P_2"*) determined from n independent repeated observations $X_{i,k}$ (i.e., results of the questionnaires on the Integrity of P_2 answered by P_1) the estimate is $x_i = \bar{X}_i$, and the standard uncertainty of the estimate x_i is $u(x_i) = s(\bar{X}_i)$, with $s(\bar{X}_i)$ calculated according to Eq. (3). For convenience, $u^2(x_i) = s^2(\bar{X}_i)$ and $u(x_i) = s(\bar{X}_i)$ are sometimes called Type A variance and Type A standard uncertainty, respectively.

However, for an estimate x_i of an input quantity X_i that has not been obtained from repeated observations (i.e., type B evaluation), the associated estimated variance $u^2(x_i)$ or the standard uncertainty $u(x_i)$ cannot be evaluated through Eq. (2) and Eq. (3). They should be evaluated *ad hoc* based on all of the available information on the possible variability of X_i (JCGM 2008). For convenience, $u^2(x_i)$ and $u(x_i)$ evaluated in this way are sometimes called Type B variance and Type B standard uncertainty, respectively. Knowledge about X_i that can help to estimate its uncertainty of measurement under type B evaluation can be, for example: "Based on the available information, one can state that there is a fifty-fifty chance that the value of the input quantity X_i lies in the interval a_- to a_+", or "Based on the available information, one can state that there is about a two out of three chance that the value of X_i lies in the interval a_- to a_+". In other cases, it may be possible to estimate only bounds (upper and lower limits) for X_i, in particular, to state that "the probability that the value of X_i lies within the interval a_- to a_+ for all practical purposes is equal to one and the probability that X_i lies outside this interval is essentially zero". If there is no specific knowledge about the possible values of X_i within the interval, one can only assume that it is equally probable for X_i to lie anywhere within it. Then x_i, the expected value of X_i, is the midpoint of the interval, $(a_- + a_+)/2$, with associated variance $u^2(x_i) = (a_+ - a_-)^2/12$. This is the case of social trust factors with a value directly provided by the decision maker. The only information provided is a value within reference limits. In this case, the first step of the calculation process must be addressed by a Type B evaluation.

Note that, in the presence of historical data, until now all the observations are taken into account in the calculation process, no matter how old they are. Some researchers consider that the age of the data is important. Thus, for example, in (Sabater and Sierra 2001) a weighted mean is used, giving more relevance to recent impressions. Xiong and Lui (2003) state that in a business community, one may wish to use the recent history of a peer and at the same time consider the historical ratings a peer received in the past but with less weight than the recent history in order to evaluate the peer based on its consistent behaviour. As another example, Ruohomaa and Kutvonen (2010) propose the definition of *epochs* in order to discount old information in favour of new through different weights.

Taking into account the previous, if one wishes to consider in this procedure the age of the observations several options arise. For example, the actual input value of each $X_{i,k}$ could reflect its age by a function, or X_i must be calculated as a weighted mean. Note that in the latter case Eq. (1) and Eq. (2) have to be reformulated as Eq. (1.a) and Eq. (2.a), i.e., weighted mean and weighted variance. The rest of the process does not change.

$$\overline{q_w} = \frac{1}{\sum_{i=1}^{n} w_k} \sum_{k=1}^{n} w_k q_k \quad (1.a)$$

$$s_w{}^2 = s^2/b \text{ where } b = (\sum_{k=1}^{n} w_k)^2 / \sum_{k=1}^{n} w_k{}^2 \quad (2.a)$$

3.1.2 Step 2: Combined uncertainty calculation

When X_i is a single variable (i.e., it is not a combination of other variables, as with variables defined in this paper at *peer-to-peer* level), $u_c(x_i) = u(x_i)$.

3.1.3 Step 3: Expanded uncertainty calculation

Having obtained the value of $u_c(x_i)$—$u_c(x_i) = u(x_i)$ in this case—either by type A or type B evaluation, it is necessary to know $U(x_i)$ since what is most important is to calculate the interval that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to X_i. In fact, Johnson and Grayson (2005) argue that cognitive trust arises from accumulated knowledge that allows the trustor to make decisions related to trustee's trustworthiness with some level of confidence. This level of confidence is provided by the confidence interval.

In this respect, as stated above, $U(x_i) = k * u_c(x_i)$. Now what remains is to know the value of k. To this end, the sampling distribution of the characteristic of interest must be estimated.

In this paper, it is assumed that the populations are normal. The normal distribution is the most widely used family of distributions in statistics and many statistical tests are based on this assumption. In fact, the recommendation in (NASA 2010) is that "the normal distribution should be applied as the default distribution, unless information to the contrary is available". Moreover, with regard to the social trust factors measured in this paper, it is known that most personality dimensions are normally distributed, so it is also assumed (Matthews et al. 2009).

With the previous considerations, a simple approach is often adequate in measurement situations where the probability distribution characterized by the estimate x_i of a measurand X_i and $u_c(x_i)$ is approximately normal and the effective degrees of freedom of $u_c(x_i)$ is of significant size. The degrees of freedom (denoted by v_i) are equal to $n - 1$ for a single quantity estimated by the arithmetic mean of n independent observations under type A evaluation and, according to the convention, are assumed to be infinite for type B uncertainties. When that is the case, which frequently occurs in practice, one can assume that taking $k = 1$ produces an interval having a level of confidence of approximately 68.27%, that taking $k = 2$ produces an interval having a level of confidence of approximately 95%, and that taking $k = 3$ produces an interval having a level of confidence of approximately 99% (i.e., $k_{68.27} = 1$, $k_{95} = 2$, and $k_{99} = 3$). On the other hand, if the decision maker has exact knowledge about a measurand data distribution different from the normal distribution (e.g., rectangular) then the values of k produce different confidence intervals (e.g., if the probability distribution is rectangular

then, for example, k = 1 produces an interval having a level of confidence of approximately 57.7%).

To obtain a better approximation than simply using a value of k_p from the normal distribution (or if this approach is not appropriate), the calculation of an interval having a specified level of confidence requires the distribution of the variable $(x_i - X_i)/u_c(x_i)$ (JCGM 2008).

If the measurand X_i is simply a single normally distributed quantity estimated by the arithmetic mean of n independent repeated observation $X_{i,k}$ of X_i (i.e., $x_i = \overline{X_i}$) with $u_c(x_i) = u(x_i) = s(\overline{X_i})$, then the distribution of the variable $t = (x_i - X_i)/u_c(x_i)$ is the t-Student distribution with $v_i = n - 1$ degrees of freedom and with

$$p = Pr\left[x_i - t_p(v_i)u_c(x_i) \leq X_i \leq x_i + t_p(v_i)u_c(x_i)\right] \quad (4)$$

In these expressions, $Pr[\ \ ]$ means "probability of" and the t-factor $t_p(v_i)$ is the value of t for a given value of the parameter v_i such that the fraction p of the t-Student distribution is encompassed by the interval $-t_p(v_i)$ to $+t_p(v_i)$. Thus the expanded uncertainty $U(x_i) = k * u_c(x_i) = t_p(v_i) * u_c(x_i)$—also called *Up*—defines an interval $x_i - U_p$ to $x_i + U_p$, conveniently written as $x_i \pm U_p$, that may be expected to encompass a fraction p of the distribution of values that could reasonably be attributed to X_i, and p is the coverage probability or level of confidence of the interval.

The results obtained during these calculations are used to complete each item (*TF*, *C*, (*R*, *U*, *S*, *N*)) in the *peer-to-peer* data structure of each partner. Thus, for example, let *X_1 = "P_1's view about the Integrity of P_2". Then* (*TF*, *C*, (*R*, *U*, *S*, *N*)) = (*Integrity, Ø, (x_1, U(x_1), s(X_1), n*)) *is an item in the peer-to-peer* data structure of P_1 that contains the information about the Integrity of P_2 in the view of P_1.

3.2 Procedure for global level

Once the *peer-to-peer* values have been calculated, the next step is to complete the *global* level data structure (Section 2) through the combination of the corresponding *peer-to-peer* values.

When measurands at *global* level are being assessed, the structure (*P*, *TF*, *C*) is proposed to express the conditions of measurement required by GUM, where *P* and *TF* are respectively the involved partner and social trust factor, and *C* is the optional activity-related context. It allows to define variables as, for example, *Y = "Integrity of P_1 in testing activities" or Y = "Integrity of P_1"*.

3.2.1 Step 1: Standard uncertainty calculation

Measurand Y is not measured directly, but is determined from N other quantities X_1, X_2, …, X_N through a functional relationship f:

$$Y = f(X_1, X_2, \cdots, X_N) \quad (5)$$

The input quantities X_1, X_2, …, X_N upon which the output quantity Y depends may themselves be viewed as measurands and may themselves depend on other quantities. As a result, the estimate of the measurand Y, denoted by y, is obtained using input estimates x_1, x_2, …, x_N for the values of the N quantities X_1, X_2, …, X_N. Thus, the output estimate y, which is the result of the measurement, is given by

$$y = f(x_1, x_2, \cdots, x_N) \quad (6)$$

The standard uncertainty calculation does not apply for y since it is not measured directly. In order to calculate y, it is necessary to firstly calculate each x_i following the

steps presented previously, and the function *f* must be defined. After that, the process continues with the next step.

3.2.2 Step 2: Combined uncertainty calculation

Whatever the function *f* that defines *Y*, the values of x_1, x_2, …, x_N cannot be determined precisely since they refer to social trust factors, so no exact value can be assigned to *y* and uncertainties come into play.

Quantities X_1, X_2, …, X_N can be correlated or uncorrelated variables (JCGM 2008). Correlation is a relationship between two variables in which both variables move in tandem (i.e., as one variable decreases the other also decreases, or when one variable increases the other also increases). Clearly, variables X_i (i = 1...N) in this paper are uncorrelated (e.g., *Y= "Integrity of P2"* defined as the combination of variables X_1 = *"P1's view about the Integrity of P2", X_2 = "P3's view about the Integrity of P2", …, X_N = "PN+1's view about the Integrity of P2"*). In this case, the uncertainty of *y* (denoted by $u_c(y)$), where *y* is the estimate of the measurand *Y* and thus the result of the measurement, is obtained by appropriately combining the standard uncertainties of the input estimates x_1, x_2, …, x_N as follows:

$$u_c{}^2(y) = \sum_{i=1}^{N} \left(\frac{\partial f}{\partial x_i}\right)^2 u^2(x_i) \tag{7}$$

Eq. (7) expresses what is termed the Law of Propagation of Uncertainty (JCGM 2008).

In addition, following GUM, when the non-linearity of *f* is significant, higher-order terms in the Taylor series expansion must be included in the expression for $u_c^2(y)$. If the distribution of each X_i is normal, the most important of such terms are

$$\sum_{i=1}^{N}\sum_{j=1}^{N}\left[\frac{1}{2}\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)^2 + \frac{\partial f}{\partial x_i}\frac{\partial^3 f}{\partial x_i \partial x_j{}^2}\right] u^2(x_i)\, u^2\left(x_j\right) \tag{8}$$

The partial derivatives, often called sensitivity coefficients, describe how the output estimate *y* varies with changes in the values of the input estimates x_1, x_2, …, x_N. This suggests writing Eq. (7) as follows:

$$u_c{}^2(y) = \sum_{i=1}^{N} [c_i u(x_i)]^2 = \sum_{i=1}^{N} u_i{}^2(y) \tag{9}$$

where $c_i = \partial f / \partial x_i$ and $u_i(y) = |c_i| u(x_i)$

3.2.3 Step 3: Expanded uncertainty calculation

The next step is to obtain $U_p = k * u_c(y)$, which implies that *k* must be calculated.

As indicated in (JCGM 2008), if $Y = c_1X_1 + c_2X_2 + \cdots + c_NX_N$ (as can be seen in Section 1, this is the type of combinations found in literature), and all the X_i are characterized by normal distributions, then the resulting convolved distribution of *Y* will also be normal. However, even if the distributions of the X_i are not normal, the distribution of *Y* may often be approximated by a normal distribution because of the Central Limit Theorem (CLT) (Nisbet et al. 2009). A practical consequence of the CLT is that, when it can be applied (in particular, if $u_c(y)$ is not dominated by a standard uncertainty component obtained from a Type A evaluation based on just a few observations, or by a standard uncertainty component obtained from a Type B evaluation based on an assumed

rectangular distribution), a reasonable first approximation to calculating U_p that provides an interval with level of confidence p is to use for k_p one of the previously mentioned values from the normal distribution (Section 3.1.3). If a better approach is required, it is possible to approximate the distribution of the variable $(y - Y)/u_c(y)$ by a t-Student distribution with effective degrees of freedom v_{eff} obtained by the Welch-Satterthwaite formula as follows:

$$v_{eff} = \frac{{u_c}^4(y)}{\sum_{i=1}^{N} {u_i}^4(y)/v_i}, u_i(y) = c_i u(x_i) \quad (10)$$

The expanded uncertainty $U_p = k * u_c(y) = t_p(v_{eff}) * u_c(y)$ thus provides an interval $y \pm U_p$ having a level of confidence of approximately p. If v_{eff} is not an integer, which will usually be the case, either interpolate or truncate v_{eff} to the next lower integer.

In order to complete the *global* level data structure, for each partner and trust factor (and context, if it is specified), the standard deviation must be obtained. The suggestion is to calculate it by combining the standard deviation of all involved X_i as if they were a single group as follows (Langley 1971):

$$s_T = \sqrt{\left(B_T - \left({A_T}^2/N_T\right)\right)/N_T - 1} \quad (11)$$

where

$$B_T = \sum_{i=1}^{N} B_i \,, A_T = \sum_{i=1}^{N} A_i \,, N_T = \sum_{i=1}^{N} n_i$$

$$B_i = [(n_i - 1)s^2(X_i)] + {A_i}^2/n_i \,, A_i = n_i \overline{X_\iota}$$

This way, each element *(TF*, *C*, (*R*, *U*, *S*, *N)*) for each partner P_i in the *global* level data structure is completed, where *TF* is the interest trust factor, *C* is the optional activity-related context, *R* takes the value *y*, *U* refers to U_p, *S* represents the standard deviation s_T of the set of all the data from questionnaires (represented through the different involved *peer-to-peer* subsets), and, finally, *N* represents the number of involved *peer-to-peer* subsets.

4 Illustrative example

In order to illustrate the proposed data structures and calculation method, an example is presented next. In this case, *Integrity* is the interest social trust factor considered for both *peer-to-peer* and *global* levels. However, as it was said previously, the proposed approach can be applied whatever the environment and social trust factor. In the case of *Integrity*, when referring to trust dimensions it can be defined as the partner honesty and promise keeping (McKnight et al. 2002) or as the perception that the partner adheres to a set of principles that the group find acceptable (Mayer et al. 1995).

The conditions of measurement for the *peer-to-peer* calculation in this example are (*Partner*$_2$, *Partner*$_1$, Integrity, Ø). It allows to define the variable X_1 as follows:

X_1 = *"Partner*$_2$*'s view on Partner*$_1$*'s Integrity"*

In this case, the calculations are based on n = 12 observations obtained from the results of the integrity questionnaire that *Partner*$_2$ has fulfilled after each collaboration with *Partner*$_1$. This questionnaire contains 14 questions addressing *Partner*$_2$*'s* view on Partner$_1$*'s* integrity. Examples of such questions are: *"Does your partner respect the existing working standards in the project/task?"*, *"Does your partner respect the existing contracts and agreements?"*, and *"Does your partner provide you with the information you need to take preventive and/or corrective actions?"*. The combination of the

responses to these questions ranges, as in (Lavrac et al. 2007), from 1 to 6. Specifically, the results obtained from the answered integrity questionnaires are:

$X_{1,j}$ = 3, 4, 3, 4, 5, 4, 5, 4, 5, 4, 5, 6; j = 1..12.

This way, the tuple (Integrity, Ø, (*R*, *U*, *S*, 12)) in the *peer-to-peer* data structure of $Partner_2$ will be progressively completed through the proposed calculations to finally obtain the opinion of $Partner_2$ about the integrity of $Partner_1$.

The first step is to calculate the estimate of the value of X_1, which is the value of *R* in the aforementioned tuple. Thus, by Eq. (1):

$$x_1 = \overline{X_1} = 4.33$$

(Integrity, Ø, (4.33, *U*, *S*, 12))

The value of *S* can be also calculated by the positive square root of the experimental variance by Eq. (2):

$$s(X_1) = \sqrt{s^2(X_1)} = \sqrt{0.787} = 0.887$$

(Integrity, Ø, (4.33, *U*, 0.887, 12))

What is still to be known is *U*, that is to say, the uncertainty of measurement. To do this, $u(x_1)$, $u_c(x_1)$, and $U(x_1)$ must be calculated. The value of $u(x_1)$ is calculated by Eq. (3):

$$u(x_1) = s(\overline{X_1}) = 0.887/\sqrt{12} = 0.256$$

X_1 is a single variable (i.e., it is not created from other variables), so that

$$u_c(x_1) = u(x_1), U(x_1) = k * u_c(x_1) = k * u(x_1)$$

In this example, the t-Student distribution of the variable $t = (x_1 - X_1)/u_c(x_1) = (x_1 - X_1)/u(x_1)$ with $v_1 = 12 - 1$ degrees of freedom will be used to calculate *k*. Thus, by Eq. (4):

$$p = Pr\left[x_1 - t_p(v_1)u(x_1) \leq X_1 \leq x_1 + t_p(v_1)u(x_1)\right]$$

Let us suppose that a level of confidence of the interval of 95% (p = 95) is enough in the CN domain. In the case of X_1, the t-Student distribution table indicates that

$$t_p(v_1) = t_p(n-1) = t_{95}(11) = 2.2$$

That is to say, $k = t_{95}(11) = 2.2$, thus the following by Eq. (4):

$$95\% = Pr[4.33 - 2.2 * 0.256 \leq X_1 \leq 4.33 + 2.2 * 0.256] = Pr[3.77 \leq X_1 \leq 4.89]$$

$$U(x_1) = U_{95} = 0.56$$

$$X_1 = 4.33 \pm 0.56$$

(Integrity, Ø, (4.33, 0.56, 0.887, 12))

Thus, the values that could reasonably (with p = 95%) be attributed to the integrity of $Partner_1$ in the collaborations with $Partner_2$ are 4.33±0.56, with a standard deviation of 0.887 in a sample of 12 observations, which results in the following tuple in the *peer-to-peer* data structure of $Partner_2$:

$Partner_1$: [(*Integrity,* Ø, (4.33, 0.56, 0.887, 12))]

Therefore, the interpretation of the tuple is the following: the value of integrity that $Partner_2$ assigns to $Partner_1$ based on her/his experience (i.e., the value of X_1) is 4.33; however, considering the uncertainty of measurement, the real value of X_1 will be one in the interval [3.77, 4.89] with a probability of 95%.

The variable X_1 defined at *peer-to-peer* level represents $Partner_2$*'s view on* $Partner_1$*'s integrity. However, if* one wants to know the integrity of $Partner_1$ then all *peer-to-peer* views on $Partner_1$'s integrity must be combined in order to obtain the *global* level value. The conditions of measurement for the *global* level calculation in this example are ($Partner_1$, Integrity, Ø). It allows to define *Y = "Integrity of* $Partner_1$*"*. In this case, as in most cases, the measurand *Y* is not measured directly, but is determined from *N* other quantities $X_1, X_2, ..., X_N$ through a functional relationship *f*. Let X_i = *"*$Partner_{i+1}$*'s view on*

Partner1's Integrity", i = 1..7, and suppose the X_i, x_i, $u(x_i)$ and $s(X_i)$ values in Table 1 (n_i = 12 and v_i = 11 in all cases):

Table 1 *Peer-to-peer* values for *Partner1*

X_i	Description	x_i	$u(x_i)$	$s(X_i)$
X_1	*Partner2's view on Partner1's Integrity*	4.33	0.256	0.887
X_2	*Partner3's view on Partner1's Integrity*	3.83	0.32	1.107
X_3	*Partner4's view on Partner1's Integrity*	4.99	0.3	1.038
X_4	*Partner5's view on Partner1's Integrity*	2.5	0.25	0.865
X_5	*Partner6's view on Partner1's Integrity*	4.41	0.3	1.038
X_6	*Partner7's view on Partner1's Integrity*	3.25	0.28	0.968
X_7	*Partner8's view on Partner1's Integrity*	4.33	0.29	1.003

Following Eq. (5):

$Y = f(X_1, X_2, X_3, X_4, X_5, X_6, X_7)$

As trust factors are usually combined in literature using average functions (Section 1), let us consider the following functional relationship:

$$Y = \frac{1}{7}\sum_{i=1}^{7} X_i$$

Thus, *Y* is the mean of the opinions on the integrity of *Partner1* of all the partners that have collaborated with her/him.

Note that if the arithmetic mean is used it is assumed that the opinion of each partner accounts equally (i.e., *1/7*) and that it is independent of the sample size. If, for example, elements such as partners' credibility/reliability, that is a subject of study in literature (e.g., Zacharia and Maes 2000; Sabater and Sierra 2001; Xiong and Liu 2003), or sample size need to be taken into consideration, each X_i must be weighted. For example, if the number of observations used to calculate each x_i (i.e., the sample size) is taken into account, then

$$Y = \sum_{i=1}^{7} X_i n_i \Big/ \sum_{i=1}^{7} n_i$$

Continuing with the example, and following Eq. (6):

$$y = \frac{1}{7}\sum_{i=1}^{7} x_i$$

$y = 3.95$

(Integrity, Ø, (3.95, *U*, *S*, 7))

Eq. (9) derives:

$c_i = \partial f / \partial x_i = 1/7$

$u_c{}^2(y) = 0.0117$

$u_c(y) = 0.1081$

In order to approximate the distribution of the variable $(y - Y)/u_c(y)$ by a t-Student distribution and calculate $U(y)$, v_{eff} must be obtained by Eq. (10) resulting in v_{eff} = 75. The t-Student distribution table indicates that $t_p(v_{eff}) = t_{95}(75) = 1.9921$. That is to say, $k = t_{95}(75) = 1.9921$ and, thus, by Eq. (4):

$$95\% = Pr[3.95 - 1.9921 * 0.1081 \leq Y \leq 3.95 - 1.9921 * 0.1081]$$
$$= Pr[3.74 \leq Y \leq 4.16]$$

$U(y) = U_{95} = 0.21$

$Y = 3.95 \pm 0.21$

(Integrity, Ø, (3.95, 0.21, *S*, 7))

All that remains to be known is *S*. In this case, as *Y* is a derived variable (i.e., defined by a function or expression in terms of other variables), the standard deviation is obtained by Eq. (11):

$S = s_T = 1.23$

(Integrity, Ø, (3.95, 0.21, 1.23, 7))

Thus, the values that could reasonably (with $p = 95\%$) be attributed to the integrity of *Partner1* are 3.95±0.21, with a standard deviation of 1.23 in a sample of 84 observations (7 partners and 12 observations per partner), which results in the following data structure:

Partner1: [(Integrity, Ø, (3.95, 0.21, 1.23, 7))]

Therefore, the interpretation of the tuple is the following: the value of the integrity of *Partner1* based on all the available opinions is 3.95; however, considering the uncertainty of measurement, the real value of *Y* will be one in the interval [3.74, 4.16] with a probability of 95%.

5 Conclusions and future work

This paper has proposed a metrology based method for the measurement of the social dimension of (cognitive) trust factors in CNs capturing the different views of trust factors in the literature. Thus, two different but related perspectives on trust among partners in CNs have been defined for each social trust factor: (i) *peer-to-peer* level (i.e., trust of one partner to another partner), and (ii) *global* level (i.e., the general trustworthiness of a partner). The *peer-to-peer* values are calculated through the combination of questionnaire results whereas the *global* values are calculated through the combination of the corresponding *peer-to-peer* values.

The value of each trust factor at both data levels is not represented through a single value but through the data structure (*TF*, *C*, (*R*, *U*, *S*, *N*)), where *TF* is the trust factor, *C* is the optional activity-related context, *R* is the most likely outcome (i.e., the estimate), *U* is the uncertainty of measurement, *S* is the standard deviation, and *N* is the number of observations. Thus, the value of *S* characterizes the variability of the *N* observed values, and *U* characterizes the dispersion of the quantity values that could reasonably be attributed to *TF,* with *R* being the most likely outcome. This structure provides a better understanding of the measurement result since it puts together information items identified in literature and new essential information about the result not provided by other approaches. This new essential information refers to the uncertainty of measurement, which characterizes the lack of exact knowledge of the value of the measurand. The subjectivity inherent in social trust indicators makes it necessary to provide this quantitative indication of the uncertainty of the result, in order to allow users to assess its suitability. This, and all data in the data structures defined at both *peer-to-peer* and *global* level can be obtained through the Metrology based calculation method proposed in this paper.

Thus, when configuring a CN, it will no longer be necessary to attend only to exact (i.e. precise and unique) values for the considered trust factors (e.g., "Empathy > 3" vs "Empathy > $3 \pm x$ based on at least 12 opinions with a dispersion of values smaller that *y*", both for data at *peer-to-peer* and *global* levels). This approach, that allows high flexibility on partner selection in CNs, entails the need to develop/adapt the systems in

the CN domain so they can adequately manage the new information presented in this paper (Section 2) and in the specified way (Section 3).

In addition, note that, as can be seen in Section 1, sometimes different trust factors of different types are combined in order to obtain the value of new trust indicators, or even a global and unique value for the trustworthiness of each partner (e.g., Tian and Wang 2012). Although this paper refers to subjective trust factors, trust indicators measured via objective methods could also be incorporated to these calculations by simply considering their uncertainty of measurement to be zero in absence of information about it.

As far as the type of involved variables, when combinations involve uncorrelated variables the method proposed in Section 3.2 to calculate the *global* level values is directly applicable. In the case of correlated variables, the first step is to study if such correlation is actually possible in the considered domain. In fact, the authors are currently analysing, for example, the available data from a web application for the creation of CNs (that implements the proposal presented in this paper as a support for the negotiation process) in order to determine if this phenomenon of correlation arises. If it finally occurs, the formulas in Section 3.2 should be adapted to this situation following the mathematical guidelines proposed by GUM for correlated variables.

The authors are also working on three additional related research issues:

1. Discovering temporal patterns and trends. In a CN, trust factors might be considered as time series, as they derive from a set of regular time-ordered observations of a quantitative characteristic of partners (e.g., integrity) taken at successive points of time. Finding temporal association patterns and trends will provide a better understanding of the trust factors involved in collaborative work, allowing to define and implement strategies for improving the success of CNs. To that end, the proposals of Aljawarneh et al. (2017a, 2017b) and Radhakrishna et al. (2017) will be taken as the starting point of the research.
2. Outlying questionnaires results when partners are expressing their opinion about others. Any suspected result must be analysed in order to decide if it must be rejected or not and how to manage the data.
3. Credibility/reliability and how to incorporate it to the proposed approach. Suspicious cases like, for example, very different *peer-to-peer* values with the same conditions of measurement, very different opinions about a given partner, or outlying results, among other indicators, could provide credibility/reliability indicators.

References

Afsarmanesh H, Camarinha-Matos LM, Msanjila SS. (2011) Models, methodologies, and tools supporting establishment and management of second-generation VBEs. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews 41(5): 692–710. https://doi.org/10.1109/TSMCC.2010.2076326.

Aljawarneh SA, Vangipuram R, Puligadda VK, Vinjamuri J (2017a) G-SPAMINE: An approach to discover temporal association patterns and trends in internet of things. Future Generation Computer Systems 74: 430-443. https://doi.org/10.1016/j.future.2017.01.013.

Aljawarneh SA, Elkobaisi MR, Maatuk AM (2017b) A new agent approach for recognizing research trends in wearable systems. Computers & Electrical Engineering

61: 275-286. https://doi.org/10.1016/j.compeleceng.2016.12.003.

Andrade J, Ares J, Garcia R, Martinez MA, Pazos J, Suarez S (2015) A Game Theory Based Approach for Building Holonic Virtual Enterprises. IEEE Transactions on Systems, Man, and Cybernetics: Systems 45(2): 291–302. https://doi.org/10.1109/TSMC.2014.2341220.

Ashtiani M, Abdollahi-Azgomi M (2015) A multi-criteria decision-making formulation of trust using fuzzy analytic hierarchy process. Cognition, Technology and Work 17(4): 465-488. https://doi.org/ 10.1007/s10111-014-0310-2.

Barki H, Robert J, Dulipovici A (2015) Reconceptualizing trust: A non-linear Boolean model. Information and Management 52(4): 483–495. https://doi.org/10.1016/j.im.2015.02.001.

Bell S (2001) A beginner's guide to uncertainty of measurement. https://www.dit.ie/media/physics/documents/GPG11.pdf. Accessed 16 March 2018.

Berry GR (2011) A cross-disciplinary literature review: Examining trust on virtual teams. Performance Improvement Quarterly 24(3): 9–28. https://doi.org/10.1002/piq.20116.

Camarinha-Matos LM (2014) Collaborative Networks in Industry and the Role of PRO-VE. International Journal of Production Management and Engineering 2(2): 53–56. https://doi.org/10.4995/ijpme.2014.3031.

Camarinha-Matos LM, Afsarmanesh H (1999) The virtual enterprise concept. In: Camarinha-Matos LM, Afsarmanesh H (eds), Infrastructures for Virtual Enterprises: Networking Industrial Enterprises. New York, Kluwer Academic Publisher, pp. 3–14. https://doi.org/10.1007/978-0-387-35577-1_1.

Camarinha-Matos LM, Afsarmanesh H (2005) Collaborative networks: a new scientific discipline. Journal of Intelligent Manufacturing 16(4): 439–452. https://doi.org/10.1007/s10845-005-1656-3.

Chen YH, Lin TP, Yen DC (2015) How to facilitate inter-organizational knowledge sharing: The impact of trust. Information & Management 51(5): 568–578. https://doi.org/10.1016/j.im.2014.03.007.

Dirks, KT (1999) The effects of interpersonal trust on work group performance. The Journal of Applied Psychology 84(3): 445-455. https://doi.org/10.1037/0021-9010.84.3.445.

Dunn, P (2000) The Importance of Consistency in Establishing Cognitive-based Trust: A Laboratory Experiment. Teaching Business Ethics 4(3): 285–306. https://doi.org/10.1023/A:1009870417073.

Fan ZP, Feng B, Suo WL (2009) A fuzzy linguistic method for evaluating collaboration satisfaction of NPD team using mutual-evaluation information. International Journal of Product Economics 122(2): 547–557. https://doi.org/10.1016/j.ijpe.2009.05.018.

Fan ZP, Suo WL, Feng B, Liu Y (2011) Trust estimation in a virtual team: A decision

support method. Expert Systems with Applications 38(8): 10240–10251. https://doi.org/10.1016/j.eswa.2011.02.060.

Fang Y, Kwok RC-W, Schroeder A (2014) Knowledge processes in virtual teams: consolidating the evidence. Behaviour & Information Technology 33(5): 486-501. https://doi.org/10.1080/0144929X.2012.719033.

Farrington-Darby T, Wilson JR (2009) Understanding social interactions in complex work: a video ethnography. Cognition, Technology and Work 11(1): 1-15. http://doi.org/10.1007/s10111-008-0118-z.

Gambetta D (1988) Can We Trust Trust? In: Gambetta D (ed.), Trust: Making and Breaking Cooperative Relations. Basil Blackwell, Oxford, pp. 213–237.

Garcia E, Giret A, Botti V (2016) Designing normative open virtual enterprises. Enterprise Information Systems 10(3): 303–324. https://doi.org/10.1080/17517575.2015.1036927.

Gefen D, Benbasat I, Pavlou PA (2008) A Research Agenda for Trust in Online Environments. Journal of Management Information Systems 24(4): 275–286. https://doi.org/10.2753/MIS0742-1222240411.

Greenberg PS, Greenberg RH, Antonucci YL (2007) Creating and sustaining trust in virtual teams. Business Horizons 50(4): 325–333. https://doi.org/10.1016/j.bushor.2007.02.005.

Haller J (2008) A Bayesian Reputation System for Virtual Organizations, Negotiation, Auctions, and Market Engineering. Lecture Notes in Business Information Processing 2: 171–178. https://doi.org/10.1007/978-3-540-77554-6_12.

Handy, C (1995) Trust and the virtual organization. Harvard Business Review 73(3): 40–50.

Hardwick J, Anderson AR, Cruickshank D (2013) Trust formation processes in innovative collaborations: Networking as knowledge building practices. European Journal of Innovation Management 16(1): 4–21. https://doi.org/10.1108/14601061311292832.

Hermoso R, Billhardt H, Ossowski S (2007) Integrating Trust in Virtual Organisations. In: Noriega, P. et al. (eds.). Coordination, Organizations, Institutions, and Norms in Agent Systems II, Lecture Notes in Computer Science 4386: 19–31, Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-74459-7_2.

Hernández-Fusilier D, Montes-y-Gómez M, Rosso P, Guzmán-Cabrera R (2015). Detecting positive and negative deceptive opinions using PU-learning. Information Processing & Management 51(4): 433–443. https://doi.org/10.1016/j.ipm.2014.11.001.

Jarvenpaa SL, Leidner DE (1999). Communication and trust in global virtual teams. Organization Science 10(6): 791–815. https://doi.org/10.1111/j.1083-6101.1998.tb00080.x.

Jarvenpaa SL, Knoll K, Leidner DE (1998). Is anybody out there? Antecedents of trust in global virtual teams. Journal of Management Information Systems 14(4): 29–64. https://doi.org/10.1080/07421222.1998.11518185.

JCGM, Joint Committee for Guides in Metrology (2008). Evaluation of measurement data — Guide to the expression of uncertainty in measurement. http://www.iso.org/sites/JCGM/GUM/JCGM100/C045315e-html/C045315e.html?csnumber=5046. Accessed 16 March 2018.

JCGM, Joint Committee for Guides in Metrology (2012). International Vocabulary of Metrology – Basic and General Concepts and Associated Terms. http://www.bipm.org/en/publications/guides/vim.html. Accessed 16 March 2018.

Kamvar SD, Schlosser MT, Garcia-Molina H (2003) The Eigentrust algorithm for reputation management in P2P networks. In: Proceedings of the 12th International Conference on World Wide Web, Budapest, May 2003, pp. 640–651. doi: 10.1145/775152.775242.

Kanawattanachai P, Yoo Y (2007) The impact of knowledge coordination on virtual team performance over time. MIS Quarterly 31(4): 783–808. https://doi.org/10.2307/25148820.

Kasper-Fuehrer EC, Ashkanasy NM (2001) Communicating trustworthiness and building trust in inter-organizational virtual organizations. Journal of Management 27(3): 235–254. https://doi.org/10.1016/S0149-2063(01)00090-3.

Krosnick JA, Presser S (2009) Question and Questionnaire Design. In: Wright JD, Marsden PV (eds), Handbook of Survey Research (2nd edition). San Diego, CA, Elsevier, pp. 263-314.

Lambrechts F, Sips K, Taillieu T, Grieten S (2009) Virtual organizations as temporary organizational networks: boundary blurring, dilemmas, career characteristics and leadership. Argumenta Oeconomica 1(22): 55–82.

Langley R (1971) Practical Statistics Simply Explained. Dover Publications, U.K.

Lavrac N, Ljubic P, Urbancic T, Papa G, Jermol M, Bollhalter S (2007) Trust Modeling for Networked Organizations Using Reputation and Collaboration Estimates. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews 37(3): 429–439. https://doi.org/10.1109/TSMCC.2006.889531.

Leina Z, Tiejun P (2012) Partner Trust Evaluation Method of Virtual Enterprise. Information Technology Journal 11(4): 524–527. https://doi.org/10.3923/itj.2012.524.527.

Matthews G, Deary IJ, Whiteman MC (2009) Personality Traits. Cambridge University Press, Cambridge.

Mayer RC, Davis HM, Schoorman FD (1995) An integrative model of organizational trust. Academy of Management Review 20(3): 709–734. https://doi.org/10.5465/AMR.1995.9508080335.

Mayer RC, Gavin MB (2005) Trust in management and performance: Who minds the shop while the employees watch the boss? Academy of Management Journal 48(5): 874–888. https://doi.org/10.5465/AMJ.2005.18803928.

Mcknight HD, Choudhury V, Kacmar C (2002) Developing and validating trust measures for e-commerce: an integrative topology. Information Systems Research 13(3): 334–359. https://doi.org/10.1287/isre.13.3.334.81.

Msanjila SS, Afsarmanesh H (2009) On development of TrustMan system assisting configuration of temporary consortiums. International Journal of Production Research 47(17): 4757–4790. https://doi.org/10.1080/00207540902847330.

Msanjila SS, Afsarmanesh H (2010) FETR: a framework to establish trust relationships among organizations in VBEs. Journal of Intelligent Manufacturing 21(3): 251–265. https://doi.org/10.1007/s10845-008-0178-1.

Nakayama MK, Binotto E, Pilla BS (2006) Trust in Virtual Teams: A Performance Indicator. In: Kumar D, Turner J (eds.). Education for the 21st Century — Impact of ICT and Digital Resources. Springer, Boston, pp. 105–113. https://doi.org/10.1007/978-0-387-34731-8_12

NASA, National Aeronautics and Space Administration (2010) Measurement Uncertainty Analysis Principles and Methods. https://standards.nasa.gov/file/2627/download?token=4bsOsYtD. Accessed 16 March 2018.

Nisbet R, Elder J, Miner G (2009) Handbook of Statistical Analysis and Data Mining Applications. Academic Press, London.

Nonose K, Okukubo A, Yoda Y, Kanno T, Furuta K (2016) Support for creating introspective reports detailing cooperative behaviors with concept maps. Cognition, Technology and Work 18(1): 71-88. http://doi.org/ 10.1007/s10111-015-0347-x.

Nonose K, Kanno T, Furuta K (2014) Effects of metacognition in cooperation on team behaviors. Cognition, Technology and Work 16(1): 349-358. http://doi.org/ 10.1007/s10111-013-026.

Panteli N, Duncan E (2004) Trust and temporary virtual teams: alternative explanations and dramaturgical relationships. Information Technology & People 17(4): 423–441. https://doi.org/10.1108/09593840410570276.

Paul DL, McDaniel RR (2004) A Field of Study of the Effect of Interpersonal Trust on Virtual Collaborative Relationship Performance. MIS Quarterly 28(2): 183–227.

Radhakrishna V, Aljawarneh SA, Kumar PV, Janaki V (2017). A novel fuzzy similarity measure and prevalence estimation approach for similarity profiled temporal association pattern mining. Future Generation Computer Systems. In press. https://doi.org/10.1016/j.future.2017.03.016.

Robert, LP, Dennis, AR, Hung, YTC (2009) Individual swift trust and knowledge-based trust in face-to-face and virtual team members. Journal of Management Information Systems 26 (2): 241-279. https://doi.org/10.2753/MIS0742-1222260210.

Ruohomaa S, Kutvonen L (2010) Trust and Distrust in Adaptive Inter-enterprise Collaboration Management. Journal of Theoretical and Applied Electronic Commerce Research 5(2): 118–136. https://doi.org/10.4067/S0718-18762010000200008.

Rusman E, van Bruggen J, Sloep PB, Valcke M, Koper R (2013) The Mind's Eye on Personal Profiles: A Cognitive Perspective on Profile Elements that Inform Initial Trustworthiness Assessments and Social Awareness in Virtual Project Teams. Computer Supported Cooperative Work 22(2-3): 159-179. https://doi.org/10.1007/s10606-012-9171-5.

Sabater J, Sierra C (2001) REGRET: reputation in gregarious societies. In: Proceedings of the 5th International Conference on Autonomous Agents, Montreal, May 2001, pp. 194-195. https://doi.org/10.1145/375735.376110.

Sabater J, Sierra C (2002) Social ReGreT, a reputation model based on social relations. ACM SIGecom Exchanges 3(1): 44–56. https://doi.org/10.1145/844331.844337.

Sako M (1992) Prices, quality and trust: Inter-firm relations in Britain and Japan. Cambridge University Press, Cambridge.

Simone A, Škoric B, Zannone N (2012) Flow-based reputation: more than just ranking. International Journal of Information Technology & Decision Making 11(3): 551-578. https://doi.org/10.1142/S0219622012500113.

Tian J, Wang Y (2012) The trust field model of partner selection in virtual enterprises. In: Proceedings of the 2012 International Conference on Artificial Intelligence and Soft Computing, Poland, April-May 2012.

Xiong L, Liu L (2003) A reputation-based trust model for peer-to-peer e-commerce communities. In: Proceedings of the 2003 IEEE International Conference on E-Commerce, Newport Beach, CA, June 2003, pp. 275–284. https://doi.org/10.1109/COEC.2003.1210262.

Xu J(D), Cenfetelli RT, Aquino K (2016) Do different kinds of trust matter? An examination of the three trusting beliefs on satisfaction and purchase behavior in the buyer–seller context. The Journal of Strategic Information Systems 25(1): 15–31. https://doi.org/10.1016/j.jsis.2015.10.004.

Yasir M, Majid A, Johnson P (2014) A methodical study of the role of trust at various development stages of virtual organisations. International Journal of Networking and Virtual Organisations 14(4): 377–387. https://doi.org/10.1504/IJNVO.2014.067891.

Zacharia G, Maes P (2000) Trust Management through Reputation Mechanisms. Applied Artificial Intelligence 14(9): 881–907. https://doi.org/10.1080/08839510050144868.