This study investigates how entrepreneurs and students associate a set of 18 concepts in order to gather their Entrepreneurial mindset. Which is the connection that entrepreneurs and students believe to be among concepts such as intuition, innovation and entrepreneur? How they think innovation is associated to risk in an entrepreneurial domain? To what extent passion is connected to the idea of entrepreneur? These questions represent some extract of concepts we asked to connect. An ad-hoc software implementing the Pathfinder algorithm produced a visual representation (simplified networks) of the mindset of each group, composed by 367 students and 29 entrepreneurs. Three questions have driven our study: (1) which are the characteristics of the entrepreneurs and students mindset? (2) Are there any differences between the two representations? (3) Are there any differences among students depending on their educational background? A qualitative inspection, supported by network centrality measures, and a quantitative analysis, based upon the number of links in common among groups' networks (closeness index) and the rank-order correlation among each couple of concepts, have shown that entrepreneurs' and students' representations differ and that these differences increase when comparing entrepreneurs with students in human and natural sciences, rather than with students in social and engineering sciences. Through additional analysis we observed that the highest differences concern concepts such as Failure, Success, Social and Regional context and, in some cases, Innovation and Risk. Suggestions for future research are presented.

I. INTRODUCTION

The investigation of the mechanisms through which people interpret and use information in order to create value and new opportunities is a central feature in entrepreneurial studies (Baron, 2002; Fauchart and Gruber, 2011; Mitchell, Busenitz, Bird, Gaglio, McMullen, Morse, & Smith, 2007; Wright and Stigliani, 2013). To elucidate how these mechanisms originate and which consequences they produce, at the behavioral level, allows scholars to develop a better understanding of the entrepreneurial action, shedding new light on the differences between entrepreneurs and non-entrepreneurs (Baron, 2004). Drawn from this assumption, the entrepreneurial cognition perspective has prompted researchers to focus on knowledge structures. According to this perspective, knowledge structures, which scholars have defined as an organized knowledge of a specific domain (Edwards, Day, Arthur and Bell, 2006), serve as a mean to understand how people assess, judge or take decisions involving opportunity evaluation, venture creation and growth (Hindle, 2004; Mitchell, Busenitz, Lant, McDougall, Morse, & Smith, 2002).

In this sense, several questions rounding the matter “How do entrepreneurs think?” need to be answered.
set of concepts that convey their *Entrepreneurial mind-set*; (2) to verify if and to what extent the students training background influences the level of similarities with entrepreneurs' knowledge structures; (3) to elucidate which set of concepts contributes to determine the highest agreement between entrepreneurs' and students' knowledge structures.

We believe with this study to bring two main contributions to the entrepreneurial literature. First, we provide new knowledge on the entrepreneurial thinking, by describing central concepts, and highlighting the logic surrounding the connections among concepts in entrepreneurs' and students' representations. Second, we estimate the differences between entrepreneurs and students by adopting both qualitative and quantitative methods. In this way we try to furnish new insights on the knowledge structures differences between entrepreneurs and students.

## II. THEORY AND HYPOTHESES

### A. Knowledge Structures and Entrepreneurship

In literature, knowledge structures are also labeled as mental models, schemas, or conceptual frameworks (Day, Arthur and Gettman, 2001; Edwards et al., 2006; Klimczok & Mohammad, 1994; Kragel, Ford and Salas, 1993). Knowledge structures or mental models try to encompass the mechanisms by which individuals describe, explain, and predict events in their environment (Mathieu, Heffner, Goodwin, Salas, & Cannon-Bowers, 2000: p. 274). Consequently, they furnish a cognitive guide for individual behavior (Davis and Yi, 2004; Bandura, 1997). It becomes clear that mental models might be a pivotal construct for entrepreneurial studies, as they serve as a base to explain why some people recognize entrepreneurial opportunities, invent new products or services, assemble resources to start and grow businesses and others do not (Baron, 2004; Greogre et al., 2011; Krueger, 2003; Mitchell et al., 2002).

In the entrepreneurship research, scholars have made reference to mental models by adopting different approaches and theoretical lenses. Some scholars, for example, have defined them through a set of cognitive factors gathering prior experiences and knowledge (Wood, McElvive e Haynie, 2013), a set of beliefs concerning growth strategies (Autere and Autio, 2000), knowledge structures (R. K. Mitchell, B.T. Mitchell, & J.R. Mitchell, 2009, for a review on key research; Patel and Fiet, 2011) and role and event schemas (Corbett and Hamideski, 2007). Autere and Autio (2000) by investigating the influence of mental models on firms growth orientation, gave evidence that previous mastery experiences, outside ownership and a managers' external reference model impact on the chosen growth oriented strategies. They drew their hypothesis from the premise that mental structures (cognitive representation of reality) represent a meta-frame which embed the managers' dominant logic of the business, helping them to deal with the complexities of the everyday business. Further, several theoretical propositions have been presented about the influence of mental models on different entrepreneurial outcomes. Kellermans and Barnett (2008), for instance, argued that mental models might impact on the recognition of environmental threats in family business, while Patel and Fiet (2011) developed a framework according to which family businesses are more effective in recognizing opportunities and in adapting their strategies to the environmental requests because family members share more easily their knowledge structures, with respect to what happens in non-family businesses. Shepherd and Krueger (2002) suggested that the team might influence individual's mental models about the perceived feasibility and desirability of entrepreneurial behavior and Lim, Busenitz, and Chidambaram (2013) proposed that greater shared mental models of the venture between founders and investors might lessen the strength of the fault-line between the two subgroups. Finally, other scholars (Dodd, 2002; Hill and Levenhagen, 1995) have focused on metaphors to investigate mental models. Metaphors, which are understood to be rudimentary mental models, have been adopted as a means to understand the cognitive processes that drive the entrepreneurial activities, by which it is possible to derive individuals’ mindset.

Operationally, different ways to measure mental models/knowledge structures exist (Van Boven and Thompson, 2003). In our study, we adopted pairwise connection of a set of generic concepts, that we described in the Methods section of this paper, as a technique to elicit the entrepreneurs' and students' *Entrepreneurial mind-set*. This technique has been largely used to assess knowledge structures in training contexts (es. Curtis and Davis, 2003; Davis et al., 2003; Day et al., 2001; Kozlowski, Gully, Brown, Sals, Smith, & Nason, 2001; Schuette, Day, McEntire, P.L Boatman, J.E. Boatman, Kawollik, and Wang, 2009; Goldsmith, Johnson and Acton, 1991) and, also, to derive team members' representations in those studies aiming to investigate the influence of shared mental models on team performances (Mohammed, Ferrand, and Hamilton, 2010, for a review). As in the mentioned studies, we adopted the Pathfinder algorithm, (Schvaneveldt, Dearholt and Durso 1988; Schvaneveldt, Durso, and Dearholt, 1989), which is a network pruning technique, implemented to derive structural representations in which the most important connections among concepts are retained. Particularly, this procedure, which relies upon pairwise connections of concepts, defines a network which includes important links as indicated by the proximity matrix (Bramagian, 1999), and produces a simplified network representation (PFNETs) of the unquired knowledge domain. We considered this knowledge structures assessment technique to be adequate for our purposes, as the PFNETs are recognized as a valid representation of individual structural knowledge (Meyer and Sugiyama,
which have proven, also, to be an effective method for comparing purposes (Schvaneveldt et al., 1989).

B. Hypotheses

Several studies recently paid a great attention to understand the differences between entrepreneurs and non-entrepreneurs (Baron and Ward, 2004). In the specific case of this study, we were interested in investigating differences in entrepreneurs’ and students’ knowledge structures. In particular, we expected the two groups to be different in the way of organizing a set of concepts through which representing the Entrepreneurial mindset. Our hypotheses rely upon the constructivism and entrepreneurial cognitive development model suggested by Krueger (2007). The author proposed that an individual moves from a relatively novice entrepreneurial mindset toward a more expert mindset, by achieving significant changes in deep cognitive structures, which are stimulated by critical developmental experiences (p. 124). By following this logic, we argued that students without previous practical experiences in venture creation and interested to become entrepreneurs have a different mindset of conceiving the Entrepreneurship idea with respect to the one of entrepreneurs, as the latter have lived practical experience that have changed their knowledge structures. Studies in entrepreneurial domain produced some results that support this hypothesis. Baron and Ensley (2006), for instance, by investigating the characteristics of “business opportunity” prototypes, showed that experienced entrepreneurs have more pragmatic and richer conceptual frameworks than novice entrepreneurs. Wood et al. (2013) noticed that prior experiences of failure affect entrepreneurs’ mental model in the way entrepreneurs evaluate new venture opportunities. They showed that who experienced prior failure, whether feel a strong fear of failure, are more conservative in reacting to positive opportunity conditions, and prove to be less enthusiastic towards new opportunities, than entrepreneurs who did not live these experiences.

In this vein, we believe that entrepreneurs who experienced venture creation and growth might have a different Entrepreneurial mindset if compared to students without previous experiences, but interested in becoming entrepreneurs. Particularly, we expected the two groups to show a different logic of connecting dots (concepts) and in weighting each of these relationships. From an operational point of view, we followed a comparison procedure implemented by Schvaneveldt et al. (1989), by which they illustrated the differences between psychologists and biologists in the way of representing a set of generic natural kind concepts, containing living thing, animal and plants. By a qualitative comparison between the derived network, obtained through the Pathfinder algorithm, they highlighted that the two groups had different representations mostly referred to the mammal concept. Further, we took into account two widely used indices that quantitatively allow to compare knowledge structures: the Closeness index and the Correlation between matrices, that are explained in detail in the Method section.

Therefore, regarding our interest of investigating the differences in the way of conceiving the Entrepreneurial mindset between entrepreneurs and students, we tested the following hypothesis:

H1. Entrepreneurs and students have different knowledge structures that represent their Entrepreneurial mindset. In particular they make different connections among concepts and they give different weights to paired relationships.

Our second purpose was to investigate if, and to what extent, students’ training background influences the level of similarity among knowledge structures. Specifically, we hypothesized that students who are more familiar with entrepreneurial topics, like students in social sciences, should share more similar knowledge structures with entrepreneurs, than students in human, natural or engineering sciences. This hypothesis is consistent with the fact that knowledge structures can adequately reflect changes due to experience and learning activities (Cope, 2003; Kraiger, Salas, & Cannon-Bowers, 1995; Krueger, 2005; Kozlowski et al., 2001) and it is in line with the multidimensional definition of learning which states that learning could be understood as a change in cognitive, affective and skill capacities (Kraiger et al., 1993). A large amount of empirical studies, for example, have demonstrated that through learning and experience, knowledge structures became more accurate and complex in representing a specific phenomenon (Curtis and Davis, 2003; Day et al., 2001; Kraiger et al., 1995; Gorman and Rentsch, 2009), and similar to those ones of an expert (Curtis and Davis, 2003; Day and Yi, 2004; Goldsmith et al., 1991; Schvaneveldt et al., 1989). Accordingly, knowledge structures could be considered as a useful operationalization of learning (Day et al., 2001; Kozlowski et al., 2001). By connecting our hypothesis to the Krueger’s model (2007), we considered that students in social sciences, through learning, have started processing information and knowledge concerned with entrepreneurial topics. It is likely that these acquisitions impacted on their knowledge organization, by making their knowledge structures more similar to the entrepreneurs’ knowledge structures.

According to the mentioned theoretical background, we tested the following hypothesis:

H2. The difference in knowledge structures between entrepreneurs and students increases depending on students’ training background. In particular, the difference between students in engineering, human and natural sciences and entrepreneurs are stronger than the differences between students in social sciences and entrepreneurs.
With an exploratory purpose we performed additional analyses in order to elucidate the sets of concepts for which entrepreneurs and sub-groups of students show the highest agreement versus disagreement.

III. METHODS

A. Sample and procedure

Students are 167, all of them coming from a single Italian University. They represent a population of students who participated to a selection for attending a six months university-based entrepreneurial-education course (Rideout and Gray, 2013), which provides, two months after the beginning, a practical sustain for start-up creation (including expert mentors, incubation in two private incubators and connection with private investors). Females are 42% and the average age is 27 (s.d. 5.60). Sub-groups of students are composed by 29 students in engineering (average age 28; s.d. 4.56; female 65%), 29 in natural sciences (average age 27; s.d. 4.65; female 48%), 45 in human sciences (average age 27; s.d. 6.44; female 26%) and 67 in social sciences (average age 26; s.d. 5.76; female 49%).

The sample of entrepreneurs is composed by 29 entrepreneurs belonging to the high-tech sector, two of them are females. As regards their educational background, they are not homogeneous. All of them are graduated and two of them have a doctoral degree. Their average age is 33 (s.d. 5.84). On average, they have 4 years of experience (s.d. 3.57), and two of them have collaborated with the national and regional government as expert entrepreneurs, by contributing to suggest practical help for accompanying the venturing of new start-ups. Other two of them have been awarded as innovators and received monetary prizes for their pitches in national competitions. All the firms are located in the same region where the University is located and all of them come from the same region.

For the students’ sample we collected all data at a time during the selection procedure, while entrepreneurs were contacted via e-mail. We selected them randomly from a list of entrepreneurs incubated in a regional private incubator and from a list of start-uppers funded by a regional program. Overall, we selected 50 entrepreneurs and received 58% e-mail responses.

The two groups completed a matrix in an excel or word format, which contained a complete list of the used concepts along with their description. From the raw data, a graph or network (with 18 nodes) was created by simply converting raw similarities [1...5] to distances [5...1] with which the weight of every edge (or link) of the graph connecting all pairs of nodes (representing concepts) was set. Data were analyzed through an ad-hoc software developed with the Python language (using Numpy and MatPlotLib) implementing the Pathfinder algorithm, all network filtering and the comparison measures.

B. Measures

1. Knowledge structures

No empirically validated procedures for selecting concepts in cognitive structures research exist (Acton et al., 1994). We followed the same procedure adopted by Van Boven and Thompson (2003) who selected concepts for the negotiation domain. Specifically, in a first step 30 concepts were extrapolated from the literature. As a second step, by means of a pilot test, involving students and entrepreneurs, the 18 concepts used to assess the knowledge structures were selected. This second step was useful to gather the set of concepts by which achieving a vision as common as possible between the researchers and the subjects under analysis. In selecting the set of concepts, we considered three factors that have received attention by researchers while studying the entrepreneurial process: (1) decision making, for which we followed the summary elaborated by Mitchell et al., (2007); (2) entrepreneurial motivation, for which we drew the concepts from the paper elaborated by Shane, Locke and Collins (2003); (3) environmental and social context (Mitchell et al. 2002; Baum and Locke, 2004; Nanda and Sorensen, 2010; Stuart and Sorensen, 2005). Table I reports the list of the used concepts. To elicit the participants’ responses, participants were asked to indicate the level of correlation between each couple of concepts, relying on a 5 point scale where 1 indicates the absence of correlation and 5 a total correlation. These values have been converted by the software to transpose similarity $s_{ij}$ between concepts $i$ and $j$ to their abstract distance: $d_{ij} = 6 - s_{ij}$. These distances are directly used to set the weights $w_{ij}$ of the links in the graphs.

2. The Pathfinder algorithm

Given any connected network with $N$ nodes $n_1, \ldots, n_N$ where the node $n_i$ is connected to any neighbour $n_j$ by a weighted link $w_{n_i,n_j}$, the Pathfinder pruning method acts as a link reduction mechanism preserving the most salient connections (Schvaneveldt et al., 1989). In practice one has to set two parameters $(r, q)$ to obtain a simplified network $PFNet(r,q)$: the $r$ parameter defines the Minkowski metric to be enforced:

$$w_{n_i,n_j} \leq \left[ \sum_{l=1}^{k-1} w_{n_i,n_{l+1}}^q \right]^{1/r} \quad \forall k = 2, 3, \ldots, q$$

while $q$ specifies that the triangle inequality must be maintained against all the alternative paths with up to $q$ links connecting any two nodes. Of course the upper bound to $q$ is $N - 1$ corresponding to a PFNet composed by the union of all possible minimum spanning trees for the network. The extreme cases of the Minkowski metric correspond respectively to the simple sum of the
weights of the single links \((r = 1, \text{ also called Manhattan distance})\) and to neglecting all weights but the highest \((r = \infty, \text{ known as Chebyshev distance})\). The second choice makes sense when dealing with a metaphor such as hydraulic systems for which a single thin pipe (a weak link) is sufficient to severely limit the flow between two reservoirs (nodes), even if the remaining pipes are huge. In the present case of a sequence of connected concepts, it seems natural to follow the hydraulic example since if any two items are judged as “disconnected”, the initial and final concepts would hardly be rationally considered to be connected regardless of the strength of the other items. When dealing with weighted networks obtained from Likert ordinal scales it is customary to set \(r = \infty\) (Schvaneveldt et al., 1989) and to determine the optimal value for the \(q\) parameter in order to prune the network by simultaneously maintaining the salient information and obtaining a stable configuration of the paired relationships. By following Schvaneveldt’s et al. procedure (1989) we computed the Spearman rank-order correlation between the raw proximity matrices and the pruned networks (averaged over the columns) for different levels of the \(q\) parameter. As shown in figure 1, for all of the groups, when \(q = 3\) the best simultaneous fit between the rankings of the derived networks and the original matrices is reached. Summarizing, we set the PFNet parameters to \(r = \infty\) and \(q = 3\) to obtain any simplified network presented in this work: the first choice means that only the weakest link (longest distance) is important when computing distances, and the second one implies that a specific path between nodes \(i\) and \(j\) can be removed if any other indirect path (between the same nodes) comprising up to 3 hops is found to be shorter. It is important to note that the network obtained from PFNet \((r = \infty, q = N - 1)\) is contained in any other PFNet with any other choice of \(r\) and \(q\).

3. **Relevance of the concepts within networks**

We borrow the centrality measures from graph theory in order to highlight the central concepts from each network: Several measures exist (Bullmore and Sporns, 2009; Morais, Olsson and Schodel, 2013; Steyvers and Tenenbaum, 2005; Watts, 2004). The simplified networks that we obtained by implementing the Pathfinder algorithm are relatively small and with a low density, thanks to the pruning procedure. Therefore, we considered adequate the two following measures: (1) Degree Centrality (DC), which represents the fraction of nodes to which a node is connected to, normalized by the total number of nodes minus one; (2) the Betweenness Centrality (BM), which is defined as the normalized fraction of times a concept should be traversed when connecting every couple of concepts in the network.

In order to investigate differences among our samples (matrices) we used the Spearman rank-order correlation and the Closeness index.

4. **Spearman rank-order correlation**

Since most quantitative results presented in this work rely on the calculation of Spearman correlations, it seems appropriate to briefly resume its definition and to discuss the differences with respect to other similarity measures used to assess the relatedness among conceptual networks. Basically it assesses how well the relationship between two variables can be described using a monotonic function. The Spearman correlation coefficient \(\rho\) is defined as the Pearson correlation coefficient between the ranked variables. For a sample of size \(n\), the \(n\) raw scores \(X_i, Y_i\) are converted to ranks \(x_i, y_i\), and it is computed as:

\[
\rho = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2} \sum (y_i - \bar{y})^2}
\]

Identical values are assigned a rank equal to the average of their positions in the ascending order of the values. The standard error of the coefficient is \(\sigma = \sqrt{\frac{12}{n(n-1)}}\) which, in the present work, being \(n = 18\), gives approximately \(\sigma = 0.15\). The value of \(\sigma\) and its functional behaviour as \(\sqrt{n}\) is comparable to our experimental results shown in figure 3 obtained by averaging over the 18 columns of the matrix. This fact explains why our results for the correlation between simple pairs of subjects (pairs of 18 x 18 matrices) is much less precise than the correlation between average matrices of larger groups. The formal procedure to compare two \(18 \times 18\) matrices is to perform 18 correlations between homologous columns (same label) and to get the average value. The column by column correlations are also kept so that a fine grained information about how much each label contributes to the average can be analyzed (see the histograms in figure 6). The Pearson correlation has shown to be only partially capable to detect subtle similarities between networks: this is due to the fact that it relies not only on a monotonic dependence of the data but also its linearity. The Spearman, on the contrary, is quite tolerant with respect to the occasional raw data offset: the Spearman between the sequences \((1.0, 3.0, 5.0)\) and \((1.0, 3.0, 3.1)\) gives 1.0 as intended, while the Pearson drops to 0.886. Finally, the Spearman correlation can be used on the raw data directly (weighted fully connected networks) without the need to prune the networks before the comparison.

5. **Closeness between networks**

One of the most used measures to estimate the similarity among networks is the closeness index (Goldsmith and Davenport, 1990). Basically it is defined as the normalized fraction of links shared between two graphs with
the same nodes. While this measure is able to estimate the superposition of two objects fairly accurately, when dealing with small graphs with few links, the number of shades of similarity is very small and the signal goes abruptly from 0 “no correlation at all” to 1 “total correlation”. This is due to the fact that this measure is performed on pruned networks (e.g. by a Pathfinder), so that only a fraction of the original raw information is still comparable. Moreover, this measure assigns the same value regardless of the weights eventually present in the networks.

6. Control variables

We controlled for gender, previous experiences of entrepreneurship, age and family engagement in entrepreneurial activities.

IV. RESULTS

A. Qualitative analysis of PFNets

Our first purpose was to describe the entrepreneurs’ and students’ knowledge structures in order to gather their mindset about the way a set of concepts related to the entrepreneurship domain is interconnected. In table I the list of concepts used to elicit knowledge structures, the centrality measures for each concept and the total number of links for each network are reported. In figure 2 the PFNets for each group are shown.

A qualitative inspection of these representations shows that in the entrepreneurs’ PFNet (see figure 2(a)) the Entrepreneur concept is the central one \([DM = .59; BM = .68]\), which gives a star shape to the network. Its main connections are with Market, Passion and Risk, as we can note by observing the thickness of the links (proportional to the weight). There are three other central concepts in the entrepreneurs’ PFNet, represented by Success, Autonomy and Risk concepts (respectively centrality measures \([DM = .24; BM = .09]; [DM = .22; BM = .10]; [DM = .24; BM = .18]\). Two loops are noteworthy: As Schvanevelt (1990) postulated, loops convey a logic of association among concepts, that allows to extrapolate more information about the stored knowledge in memory, which is invisible by only using Euclidean distance measures (Barabasi, 2003). The small loop contains the concepts Entrepreneur, Innovation and Risk, which in turn is connected back to Entrepreneur. The second
loop is larger and connects the consecutive concepts Entrepreneur, Innovation, Risk, Failure, Autonomy, Success, Earn, Market and finally, again, Entrepreneur.

In students’ PFNet, the Entrepreneur concept is the central one \([DM = 4.1; BM = .58]\), as in entrepreneurs’ PFNet. The most important connections among Entrepreneur and the other concepts are with Passion, Planning, Innovation and Market. We note that Self-efficacy and Region are the two following most connected nodes in the network (respectively centrality measures are \([DM = .29; BM = .32]; [DM = .24; BM = .08]\)). Five loops can be noted. One important characteristic of these loops is the centrality of the Self-efficacy concept, that connects two of the most important loops directly to Entrepreneur.

If we compare the two representations, two differences are noteworthy: In students’ PFNet the concept Intuition is connected to the Innovation, while in entrepreneurs’ PFNet, Intuition is strongly and directly related to Entrepreneur. Further, we observed that while entrepreneurs firmly connected the concept Risk both to Innovation and Failure, students connected Risk only to Failure.

By inspecting the PFNet of each students’ sub-group, it emerges that, overall, students’ representations have a lower degree of centrality for the concept Entrepreneur, and Self-efficacy is the concept that constitutes their second most important hub.

B. Detecting differences among correlations

In order to address the two hypotheses we adopted a uniform approach. Specifically, we compared students’ and entrepreneurs’ knowledge structures via two different indices: (1) the rank-order correlation between matrices, and (2) the closeness index. Hypothesis 1 stated that entrepreneurs and students have different knowledge structures, in particular we hypothesized that entrepreneurs knowledge structures differ from those of the students in the way of ranking among paired relationship priorities (Spearman rank-order correlation), and in the way of interconnecting concepts (closeness index).

To test both hypotheses we performed a measure of correlation within the same group (within a group or intragroup) and among different groups (between groups or intergroup).

1. Between groups

To measure the similarity between two groups of subjects, the basic idea is to take two subjects \(i\) and \(j\) at a time, one from each group, and compute the Spearman correlation \(\rho_k(M_{ik}, M_{jk})\) for each column \(k\) of their matrices \(M_i, M_j\) and the to compute the average value over all the columns:

\[
\rho(M_i, M_j) = \frac{1}{n} \sum_{k=1}^{n} \rho_k(M_{ik}, M_{jk}).
\]

The issue with this approach is that the standard deviation of \(\rho\) is of the same order of magnitude of the signal we are looking for (differences of mean correlations between pairs of groups). Of course the standard deviation decreases when doing the average correlation between all pairs of subjects in two groups, but it is still too large to significantly reject the null hypothesis (i.e. the mean correlation between groups is different from the mean correlation between two other groups).
The source of the “noise” comes from the interaction of the Likert scale (1...5) with the rank-order correlation. When people are asked to give a value for the connection between two concepts using integer numbers, they inevitably assign a value with some random fluctuation which in our case is of the order of ±1. What happens is that even people sharing similar ideas produce raw matrices with many fluctuations around some average value for
each connection. If the similarity between two columns belonging to two like-minded subjects is computed, the rank-order correlation rapidly falls with a few rank inversions due to the inherent fluctuations in the raw data.

To circumvent this difficulty we propose a solution by avoiding the direct comparison of single subject rank-orders. Instead, we perform the comparison over the average matrices of small subgroups. This averaging process leads to much smoother rank orderings which can effectively be compared without the noise induced by random rank-order inversions due to the inevitable subjects’ uncertainty when assigning a value to a link. More formally: Given two distinct groups $A$ and $B$ to be compared, each composed by $n$ subjects, from combinatorics, we can form $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ subgroups (for each group) when the size of the subgroup is $k$. Each subgroup contains $k$ subjects of which we retain just the average matrix. We obtain two derived groups $A'$ and $B'$ containing not single-subject matrices but subgroup-averaged matrices. To retain the full statistical power of the original groups, it is theoretically required to populate the derived groups with all possible subgroups given by the binomial coefficients, but in practice a small sampling is sufficient given that the raw matrices are not fully independent: We numerically checked that convergence was reached both for the mean value and for the standard error after a few hundred iterations with a simple random sampling, so that a value of 500 steps was used for all the results.

In figure 3(a) the lower 5 curves (full square, empty and full triangle) show the intergroup correlations between all the entrepreneurs and all students, social science students, engineering students, natural science students and human science students respectively. These lower 5 curves are not very well separated for low $k$, but the relative order is quite stable with varying subgroup size: the social science and engineering students (circles) are the most correlated with entrepreneurs while natural and human science students (triangles) are appreciably less correlated via the Spearman coefficient. Almost nothing can be said when comparing single subjects ($k = 1$) for these 5 curves. To start seeing some difference between the correlations of the various groups of students with the entrepreneurs, $k$ should be set at least at 10. The differences might be more pronounced if the Spearman coefficient would be averaged over a select set of concepts as one can infer from the fine grained histograms in figures 6 and 7 where the various degrees of agreement between entrepreneurs and each group of students are easily visible. In figure 3(b) we see how the standard deviation of the correlation between group pairs behaves for increasing $k = 1...14$: these values of $\sigma$ are useful to estimate the overlap among the distributions of correlations and from them we obtain the t-test statistics significance discussed later.

2. Within a group

Correlations between subjects belonging to the same group have been estimated by partitioning each group
Figure 3: Spearman Rank-order Correlation: Mean Values and $\sigma$.

(a) Entrepreneurs and Students symbols are (independent) intragroup correlations: the two groups show similar homogeneity for any subgroup size. Each mean value is obtained by repeating and averaging a Spearman correlation (500 times) between random samplings of average matrices computed for subgroup sizes from 1 to 14.

(b) Average standard deviation for intragroup (stars and void squares) and intergroup correlations performed for increasing subgroup size.
into two independent sets and then performing the procedure outlined in the paragraph above.

In figure 3(a) the two top symbols (stars and empty squares) stand for the within group correlations for entrepreneurs and students. Subgroup size varies from 1 (single subject comparison) to 14: For all these values the correlation within group is consistently higher than any between groups correlation. This fact is a proxy of group homogeneity. In figure 3(b) we see how the standard deviation behaves for increasing subgroup size: For small $k$, the error is of the same amplitude as the differences we are trying to detect and the significance (see the Welch's t-test in figure 4) is quite low. To detect an appreciable effect size $k$ should be set at least to 5, where the difference between the two intra group top curves and all the inter group curves are separated by roughly their standard deviation.

3. Welch's t-test

Since each group has a different number of subjects and the standard deviations are similar but not equal, a Welch-corrected significance t-test for the Spearman correlation differences between mean values was chosen. Correlations between matrices obtained by averaging over larger groups are less noisy and lead to more significant differences as expected. In figure 4 we can see how the t-value increases almost monotonically with subgroup size. Since the number of degrees of freedom is approximately given by the number of subjects in each group divided by the chosen subgroup size, checking on a one-tailed t-test significance table, p-values smaller than 0.05 are attained for subgroup sizes larger than 3 for most groups, except for the top curve which, showing a large degree of separation from the others, has a p-value of 0.1 even when doing single subject correlations ($k = 1$), rapidly reaching p-values smaller than 0.001 for $k > 5$.

To better appreciate the results, mean values and standard deviations do not tell the full story and it is interesting to observe the whole probability distributions of the standard deviations. For the sake of clarity, we removed the intragroup student correlation (almost superimposed to the entrepreneurs' intragroup correlation shown in full black) and the intergroup all students-entrepreneurs (an average of the other 4 students' curves on the left). It is clear that social science and engineering students are more likely to overlap with the entrepreneurs with respect to human and natural science as already seen from their average values of figure 3(a). All four students' groups show a long tail on the left with low correlation in similar proportion, while the right rapidly-falling tails are what makes most of the difference in the average values. This graph has been obtained from five independent simple random samplings with subgroup size $k = 13$ after accumulating statistics for 500 steps for each curve: Each point of the curves represents the relative probability to choose two groups of 13 and finding the corresponding correlation of their average matrices (average over all columns).

Table II reports the rank-order correlation and t-test results.

C. Additional analysis

To develop a better understanding of the single components which contributed to produce the average of the Spearman correlation (through which accounting the similarity within and between groups), we decomposed the Spearman coefficient by isolating the correlations for each concept. As depicted in figures 6 and 7 entrepreneurs and students are less in agreement when ranking the following concepts: Social and Regional context, Success, the Modify events to create new opportunities, Failure, Autonomy. It is worth noting that Modify events to create new opportunities, Failure, and the Regional context represent the concepts with a less agreement also among entrepreneurs. The concepts Innovation and Risk, indeed, are borderline, as shown by the sorted histograms 7(f and g). We observe an opposite trend for entrepreneurs, among which Risk and Innovation are the concepts with the highest level of agreement.

V. DISCUSSION AND CONCLUSIONS

Our first purpose was to describe the peculiarity of entrepreneurs' and students' knowledge structures about a set of concepts through which to gather their respective idea of Entrepreneurship. We used the Pathfinder algorithm to derive the networks of each group, allowing the characterizing links to be found. Visual representations were derived from the average matrix of each group, which furnished the base for computing the centrality measures as a support for the descriptive analysis. This inspection provided the starting point from which to elucidate the logic of each connection behind the representations of each sub-group. Further, it gave a first qualitative signal on the similarities and differences among entrepreneurs and every sub-group of students, that we subsequently investigated through quantitative measures. Specifically, we noticed that the Entrepreneur concept is well connected in all networks and that entrepreneurs and students differ mostly as regards the secondary hubs. In the former, Risk, Success and Autonomy are central, while in the latter it is Self-efficacy.
Correlations between matrices obtained by averaging over larger groups are less noisy and lead to more significant differences. Since the number of degrees of freedom is approximately given by the number of subjects in each group divided by the chosen subgroup size, p-values smaller than 0.05 are attained for subgroup sizes larger than 3 for most groups, except for the top curve which has a p-value of 0.1 even when doing single subject correlations.

Table II: Rank-order Correlations and t-test Results

<table>
<thead>
<tr>
<th>Group A</th>
<th>N Group B</th>
<th>N Spearman SD</th>
<th>t-test* (k=13)</th>
<th>Closeness SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entrepreneurs 29 Students</td>
<td>167</td>
<td>.68</td>
<td>.04</td>
<td>7.5</td>
</tr>
<tr>
<td>Entrepreneurs 29 Social Science</td>
<td>67</td>
<td>.69</td>
<td>.04</td>
<td>5.0</td>
</tr>
<tr>
<td>Entrepreneurs 29 Engineering</td>
<td>29</td>
<td>.69</td>
<td>.04</td>
<td>4.0</td>
</tr>
<tr>
<td>Entrepreneurs 29 Human Science</td>
<td>45</td>
<td>.67</td>
<td>.03</td>
<td>4.0</td>
</tr>
<tr>
<td>Entrepreneurs 29 Natural Science</td>
<td>29</td>
<td>.66</td>
<td>.03</td>
<td>4.0</td>
</tr>
</tbody>
</table>

* p < 0.001 for all tests

which shows to be central. By analyzing the loops among concepts, we observed some interesting findings: Entrepreneurs and students seem to have a different vision of Innovation, which the former connect to Failure besides Entrepreneur and Risk, as do the latter. Further, while Intuition is directly connected to the Entrepreneur concept in entrepreneurs’ representation, it is connected to innovation in students’ representation.

As a second goal, we aimed to study in depth if a difference can be detected as far as the weights of the pair correlations are ordered and in the main links among the concepts. The results indicate that entrepreneurs and students differ in the way of weighting the connections between concepts and the difference grows when comparing entrepreneurs and students in human and natural science. Contrary to our expectations, we did not find significant differences between entrepreneurs and students in the way of linking the concepts one another. This finding could be due to the intrinsic characteristics of the closeness index (as we have described in the Method.
Figure 5: Differences between Entrepreneurs and Student Groups

Overall, our study provides three main contributions. First, these findings indicate that, for a generic set of concepts commonly associated to the Entrepreneurship idea, entrepreneurs and students are different in the way of organizing priority among the selected concepts. This study has not the goal of investigating the influences that the correlation and closeness indices exert at the performance level. Then, we cannot estimate their behavioral consequences. This limited capacity of forecasting a consequent scenario is also due to the mixed results that previous studies have provided in investigating the effects of closeness and correlation similarity on the performance level, which have failed in clarifying the specific role played by the two indices in explaining the performance (Schuelke et al., 2009; Smith-Jentsch, Mathieu, & Kraiger, 2005). We believe that future research in the entrepreneurship domain should elucidate the role of these indices in affecting the entrepreneurial action, and the entrepreneurial learning process.

Second, we observed that the differences between entrepreneurs and students grow when students’ training background is taken into account. Specifically, we estimated the value of these differences, by noticing that students in human and natural sciences have stronger differences with entrepreneurs, more than students in engineering and social science. We consider this finding having an important implication in the domain of entrepreneurial training. Previous research has demon-
Figure 6: Concept Consistency among Groups
(a) Entrepreneurs (intragroup)  (b) Students (intragroup)
(c) Entrepreneurs and Students  (d) Entrepreneurs and Engineering Students

It could be useful to assess knowledge structures before starting a training course in order to better predict the outcomes that students will achieve through the training activities, and to better understand the learning processes involved in reaching them.

Third, we shed light on the concepts for which there is high or low consistency between entrepreneurs and students. According to our findings entrepreneurs and students are less consistent in representing Failure, Innovation, Risk and Events (the possibility of modifying events to create new opportunities) and in the role played by the Environmental and the Social context. This result
gives support to previous inquiries which demonstrated that experienced and novice entrepreneurs have different focuses. The former are more aware about danger and opportunity, while the latter are more attracted by newness and novelty (Baron and Ensley, 2006; Baron and Shane, 2008). Obviously, we cannot draw any final conclusion with this study as regards differences between expert and novices. However, we believe that this study provides some qualitative support to test similar hypothesis in future research. Our findings suggest that new investigations are needed to develop a better understanding of those concepts that are likely to change along the entrepreneurial experience.

This study has a number of limitations that future research should consider to ensure the robustness of the results. Specifically, the research design restricts our
understanding about the role played by the experience in generating entrepreneurs’ knowledge structures. Although we noticed differences between entrepreneurs and students interested in becoming entrepreneurs, we believe new research should be conducted in order to confirm that these measured differences between them have been generated by the experiential process and not by not-controlled characteristics of our entrepreneurial sample. Particularly, future research should carefully address the description of knowledge structures by adopting a longitudinal perspective. This implies having to observe and record knowledge structures through the entrepreneurial process phases, from the opportunity recognition to the harvesting rewards (Baron and Shane, 2008). The second limitation concerns the set of concepts we used to elicit knowledge structures. In future research we believe that differences between entrepreneurs and students should be investigated with different sets of concepts. Specifically, we consider useful to record correlation and closeness similarity by using different sets of concepts in which the technicality is gradually incremented, by highlighting in which conditions the differences occur. This research design should ensure a greater understanding of the variability of the phenomenon under the different conditions of knowledge domain.

We have proposed the inspection of knowledge structures in the entrepreneurship domain as a procedure to better understand the reasoning and beliefs surrounding the entrepreneurial action. We are also encouraged to take this direction in the light of recent results in neuroscience. It does not directly apply to this contribution, but it is noteworthy that neuroscientists working on neuroplasticity observed a change in adult humans’ brain structures after a three-months learning section dealing with cognitive skills (Driemeyer, Boeke, Gaser, Büchel and May, 2008). It is not news that the loop between brain structure and brain function are a basis for cognition and learning (Zatorre, Fields, and Johansen-Berg, 2012), but an experimental demonstration of this loop still represents a complex challenge. However, thanks also to the neuro imagining technology, several studies have successfully started revealing anatomical group differences reflecting skills, knowledge or expertise (es. Bengtsson, Nagy, Skare, Forssman, Fossberg, & Ullen, 2005; Draganski, Gaser, Busch, Schulerer, Bogdahn, & May, 2004; Maguire, Gadian, Johnsrude, Good, Ashburner, Frackowiak, & Frith, 2000). Overall, these results and the entrepreneurial cognition perspective indicate the inspection of knowledge structures as timely. To intercept and understand the semantic and structural variation of the organization of the knowledge, over the entrepreneurial experience should be a challenge for future studies. This paper is a first attempt to furnish a contribution towards this direction.

Acknowledgments

We would like to thank Gianluigi Zanetti and Giovanni Busonera for useful discussion.

References


Bullmore, E., & Sporns, O. 2009. Complex brain networks: graph theoretical analysis of structural and


