Understanding the heterogeneity of corporate entrepreneurship programs

Selig, C.J.

3 Research method

The aim of this study is to deepen the understanding of (a) the organizational design of CE programs and (b) the value creation of CE programs. For both, there is a high degree of uncertainty, which can be mostly traced back to the emergence of new organizational designs and new fields of application whereas both still lack empirical investigations (see, e.g., Kurpjuweit & Wagner, 2020; Selig et al., 2018; Hill & Georgoulas, 2016). To understand the full heterogeneity of CE programs, a comparatively large dataset consisting of 67 semi-structured interviews representing 54 cases was created.

Qualitative research, by its nature, focuses on improving the understanding of a practical phenomenon and is well established in the fields of strategic management and CE (see, e.g., Shankar & Shepherd, 2019; Weiblen & Chesbrough, 2015; Nag & Gioia, 2012; Davis & Eisenhardt, 2011; Kuratko et al., 2001). A qualitative research approach seems well suited for studying the heterogeneity of CE programs for at least two reasons. First, new organizational designs are currently emerging in practice that has not yet been studied. Defining and distinguishing them from existing ones may require different and perhaps previously unknown (or unused) design elements. Second, it is unclear what types of outputs can be created by CE programs, as no harmonized set of outputs exists to date. Consequently, there are no previous studies that focus on comparing outputs across CE programs.

This chapter describes the research method and data to illustrate the empirical basis of our study. Section 3.1 will focus on the data collection. In Section 3.2, a description of the data set is given. Finally, in Section 3.3, our data analysis will be presented.

3.1 Data collection

For data collection, we will deal with the following topics. (A) semi-structured interviews, (B) the content of the interviews, (C) the composition of the data set, (D) description of the roles of the interviewees, and (E) sampling strategy.

A: Semi-structured interviews

We collected semi-structured interviews over a period of five years, starting in June 2015 and ending in July 2020. Most of the interviews were combined with an on-site visit of the respective CE program, while mainly in the last months of the data collection, the interviews

were conducted via telephone or video conferencing tools (due to the corona crisis). The use of semi-structured interviews with mostly open-ended questions made it possible to uncover and follow up on new topics that emerged during the interviews and to adapt to the specific characteristics of the respective CE program.

B: Content of the interviews – Interview guideline

The conducted interviews focus on five key themes described in Table 3.1. The interview guideline with all questions can be found in Appendix 1.

Key themes	Focus of investigation	Sample questions
(1) Personal background	Personal information	- What is your name/age? - Educational background?
	Professional career	- How did your professional career develop until today?
	Current job position	- Can you describe your current role within the CE program?
(2) CE program background	Reason for initiation	- What was the company's motivation for starting the CE program?
	Mission and objectives	- What are your objectives? - How are they measured?
	Development over time	- Did you have major adjustments in structure and/or strategy?
(3) Organization of CE program	Structural embedding	- Where are you located in the companies' organigram?
	Program structure	- Are you organized as an own legal entity?
	Level of autonomy	- Which processes/rules are different from core organization?
	Steering of CE program	- How are you controlled/steered by the management?
(4) Mode of operation	Scope of activities	- What are the tasks/activities to run the CE program?
	Program workflow	- Can you explain the process of the program and the stages?
	Type of support	- What do you offer to support the innovation projects?
(5) Value creation and impact	Project results	- How many new businesses or products did you develop?
	Spill-over effects	- What processes, etc. did change in the core organization?
	Success and impact	- What is a successful project in your understanding?
	Enabler/disabler	- What factors were critical for being successful?

Table 3.1: Structure and content of the interviews

C: Composition of the data set

For the data collection, 59 different persons were interviewed. In order to collect missing information, eight follow-up interviews with the same person were conducted, leading to a total number of 67 semi-structured interviews (according to the counting of our data set). These interviews refer to 54 cases belonging to 36 different companies (see Figure 3.1).

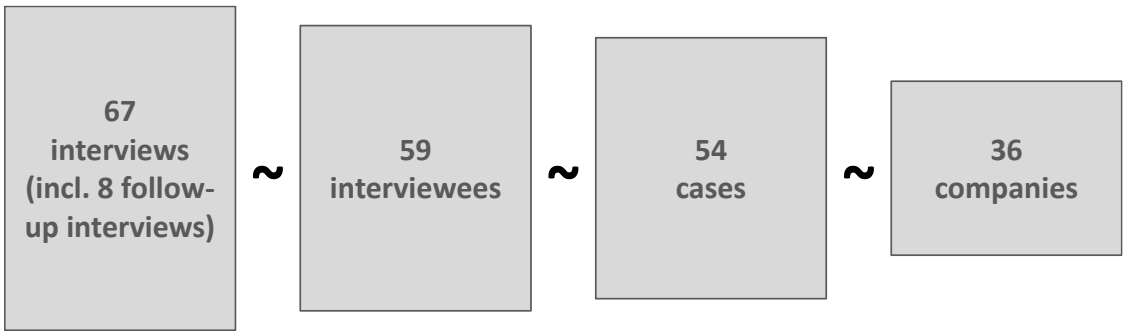

Figure 3.1: Composition of the data set

The discrepancy between the 59 different interviewees and the 54 cases can be explained by the circumstance that five people were interviewed about CE programs that were already included in our dataset. This was done to obtain additional information that could not be provided by the first interviewee. The reasons for this were either that (a) the interviewee entered the CE program at a later stage and did not know about certain facts or (b) the interviewee did not have access to the information due to his or her job position.

The distribution of these 54 cases among the 36 companies is as follows.

- 24 companies with one CE program = 24 cases
- 9 companies with two CE programs = 18 cases
- 1 company with three CE programs = 3 cases
- 1 company with four CE programs = 4 cases
- 1 company with five CE programs = 5 cases

The numbers listed above represent the interviews conducted during our research. This does not reflect the actual number of CE programs operated by the company. For example, in the case of the company where we interviewed five CE programs, there are two others that could not be interviewed due to limited access or changes in the CE programs' leadership team.

D: Description of the roles of the interviewees

The following step is the description of the four different job roles of the 59 interviewees. The number of interviewees per role is indicated in parentheses (# interviewees + # follow-up interviews) after the role name.

1) *Program lead* (40 + 4): The program lead is the person who is responsible for managing the CE program. It typically includes the design and implementation of the program as well as the management of the stakeholders from the core organization.
2) *Managing director* (4 + 1): The managing director is leading a program that is investing in startups. Due to the investments, these programs are quite capital intensive and have different structures, so they can act similarly to professional investors.
3) *Program specialist* (12 + 2): The program specialist can play different roles in the field of business development, viz. marketing or method expert. The program specialist is responsible for the support of the innovation ideas within the program.
4) *Investment manager* (3 + 1): The investment manager is supporting the investment process in startups by scouting and evaluating startups. The investment manager acts as the point of contact for the portfolio companies.

In addition to the semi-structured interviews, further data was collected (see Appendix 2) to ensure triangulation of the information (Yin, 2013). The additional data consists mainly of field notes, social media activities, scientific and practical publications, presentations, press releases, and the website of the respective CE program and the core organization.

E: Sampling strategies

For the data collection, we applied two different sampling strategies, namely *snowball sampling* and *purposive sampling*. They were needed to address the respective situation adequately. Based on two sampling strategies, we can split our data collection into two phases, which will be described below.

Phase 1 (2015 – 2017): In the first phase, we applied a *snowball sampling strategy* for our data collection. This sampling strategy is based on recommendations from one interviewee to another, which allows identifying study participants who are rather difficult to identify. When we started to conduct our semi-structured interviews in June 2015, CE was still a novel topic,

and the community in the DACH[5] region was still quite young as not many companies were operating CE programs (only eleven out of 54 cases from our data set did exist before 2015). So, we decided to focus on building a network by using peer-to-peer recommendations from one interviewee to another, which resulted in a total number of twelve interviews (from twelve different cases) that were conducted in phase 1. Based on these twelve cases, the scope of our study became clearer. In addition, CE programs were experiencing high popularity at that time, which can be seen from the fact that 43% of the cases in our data set were launched in the years 2015 and 2016. Due to this changing situation and the better visibility of potential interviewees, we changed the sampling strategy around the end of 2017.

Phase 2 (2018 – 2020): In the second phase, we switched to a *purposive sampling strategy*. This means that we from then purposively selected our interview partners according to the following two aims. First, we wished to collect interviews with all different types of CE programs that were known at the time. Second, we wanted to have multiple cases for each type covered in our data set. During phase 2, a major part of the largest companies in the DACH region were operating CE programs as an element of their innovation and digitalization strategy (Lehnen et al., 2020). This high popularity and visibility allowed us to conduct an additional 55 interviews until July 2020 (end of phase 2). At the end of our data collection, we had created a data set consisting of 67 semi-structured interviews that were representing 54 cases.

To reach theoretical saturation, 15 - 30 interviews are usually recommended (see Mason, 2010). In our case, we exceed this number by a factor of two. This was done on purpose to ensure that the variety of CE programs is fully covered and at least three cases per CE program type were included in our data set.

In Subsection 3.2, background information about the data set will be provided, followed by a general description of the data analysis in Subsection 3.3.

[5] The DACH region comprises the three countries Germany (D - Germany), Austria (A) and Switzerland (CH - Confoederatio Helvetica). In these countries the German language is largely spoken as a native language. Due to geographical proximity and similarities in their national culture, only companies based in Germany, Austria and Switzerland were covered in this study to reduce the impact of cultural differences.

3.2 Data set

This section will focus on describing the data set of our study. As shown in Figure 3.1, our data set consists of 36 different companies that are representing a total number of 54 cases. We will start this section with three information items having a general background.

1) All interviews were recorded and transcribed for analysis. This resulted in 4,533 minutes of recorded material with a total of 2,257 pages. On average, an interview lasted 68 minutes, with the shortest interview lasting 32 minutes (excluding follow-up interviews) and the longest 117 minutes.
2) Except for one company, all others have their headquarters in Germany, Austria, or Switzerland, which is known as the DACH region. The one exception is a German subsidiary of a large Spanish multinational company. The subsidiary was originally a German-based company that was acquired. It is therefore expected that the subsidiary is culturally similar to the rest of the data set.
3) Eleven of the 54 cases are CE programs located in Silicon Valley. These so-called innovation outposts (Di Fiore, 2017) are initiated to identify and understand trends early on. Since Silicon Valley is recognized as one of the world's leading innovation clusters, many companies from DACH region have launched innovation outposts there.

The next subsections will describe the data set from two different perspectives, namely the company perspective (Subsection 3.2.1) and the CE program perspective (Subsection 3.2.2).

3.2.1 Data set from a company-perspective

The companies in the data set reflect the various facets and characteristics of the DACH-industrial landscape, e.g., the high number of hidden champions, the German Mittelstand[6], and world-leading multinationals. According to configuration theory, the context of an organization has an influence on its structure and performance (see Subsection 2.2.3). Consequently, our data set should not only cover a high level of heterogeneity regarding the different organizational designs of CE programs but also take into account the companies that

[6] The Mittelstand describes the small and medium-sized companies that can be considered the backbone of the German industry. In addition to size, also the ownership structure (family-owned or family-controlled), the management style (patriarchal culture, flat hierarchy & high degree of informality) and the way of thinking (longevity & value orientation) characterize the Mittelstand (see, Pahnke and Welter (2019); Schmitt et al. (2018)).

are operating them. It is expected that a heterogeneous data set provides results that are more stable, independent from the context the CE program is operating in.

Subsequently, the data set will be described by (A) the industry classes, (B) ownership structure, and (C) company size. This illustrates the level of heterogeneity of the companies covered in our data set.

A: Industry classes

We start our description with the different industry classes to which the companies from our data set belong. For assigning a company to an industry, the NAICS (North American Industry Classification System) was used. NAICS is a common standard to classify industries, e.g., used by the federal statistic agency of the United States of America (see Pierce & Schott, 2012). The codes used in NAICS show different resolution levels, starting with two-digit numbers (11-92) that are representing broad industry titles. For example, NAICS code 11 stands for "Agriculture, Forestry, Fishing and Hunting". Below this number, a more detailed classification can be found, e.g., 11 1211 for "potato farming".

A database (EBSCO host[7]) was used to identify the NAICS code for the companies. Small companies are sometimes not found using the database. Thus, it required a second manual step to assign all companies with NAICS codes. When a company is operating in multiple industries, each NAICS code was counted to show the full range of industries. As a result, more NAICS codes than companies are used to describe the data set in terms of industry classes.

Table 3.2 illustrates the distribution of industry classes in our data set. The table shows eleven different industry classes. It is ordered by percentage, starting with the industry class that shows the highest number in our data set. The largest proportion of companies (40%) is assigned to the manufacturing industry, followed by finance and insurance with 17%. The chemical and pharmaceutical manufacturing, as well as the information class, follow with 11%. Thus, our data set covers a broad spectrum of different industries.

[7] EBSCO host – is service to access various full text database of scientific publications, which is offered by the company EBSCO Information Services.

Industry class	NAICS code	% in our data set
Manufacturing	33	40%
Finance and Insurance	52	17%
Manufacturing (chemical and pharmaceutical)	32	11%
Information	51	11%
Transportation and warehousing	48+49	5%
Utilities	22	3%
Retail trade	44	3%
Professional, scientific, and technical services	54	3%
Other services (except public administration)	81	3%
Construction	23	2%
Administrative, support, waste management, & remediation services	56	2%

Table 3.2: Overall industry classes (NAICS codes)

B: Ownership structure

In our data set, the ownership structure plays an important role. For the companies in our data set the following six types of ownership were identified.

1) *Listed public company:* A company that is listed on the stock exchange and is not majority-owned by one or more families. In most cases, these companies operate as Societas Europaea (SE), which is in the DACH region defined as Aktiengesellschaft (AG).
2) *Family ownership:* A company that is substantially controlled by a family (or a limited number of owners with family ties) and is not listed on the stock exchange.
3) *Hybrid ownership:* A company listed on the stock exchange but is usually still controlled by one or more families.
4) *State-owned:* A company that is entirely or partially owned by the state. That is either Germany or Switzerland in the case of our data set.
5) *Miscellaneous:* Different legal forms of companies that are considered to be rather rare. For example, foundation companies, cooperatives, or research companies.
6) *Subsidiaries:* A special group within our data set are subsidiaries of larger companies. All cases that belong to this category are companies that appear on the market as their own brand but belong to a larger corporation or holding structure.

The subsidiaries in our data set show different types of ownership (family ownership, hybrid ownership, listed corporations, and foundation companies). Since some of them are significantly smaller than the other companies with the same ownership structure, they are highlighted separately in Table 3.3 and Table 3.4.

Within our dataset, companies are distributed in terms of ownership structure as follows.

Ownership (# of companies)	Distribution in %
1. Listed public company (7)	19.44 %
2. Family ownership (8)	22.22 %
3. Hybrid ownership (6)	16.67 %
4. State-owned (5)	13.89 %
5. Miscellaneous (5)	13.89 %
6. Subsidiaries (5)	13.89 %

Table 3.3: Division of the companies according to the type of ownership

C: Company size

For a better understanding of the companies in our data set, they are differentiated according to their size (number of employees) in Table 3.4. We use the following three different types of values to provide a good overview: (1) $E_{MIN,}$ which describes the minimum number of employees, (2) $E_{MAX,}$ which describes the maximum number of employees, and (3) $E_{AVG,}$ which describes the average number of employees.

Ownership	E_{MIN}	E_{MAX}	E_{AVG}
1. Listed public company	34,523	385,000	173,210
2. Family ownership	1,097	64,585	33.179
3. Hybrid ownership	39,000	240,000	113,517
4. State-owned	10,000	546,924	222,249
5. Miscellaneous	2,288	147,797	53,940
6. Subsidiaries	665	58,000	13,028
Total	665	546,924	104,752

Table 3.4: Division of number of employees according to the ownership structure

Following the classification of company sizes according to the common definition used in the European Union (European Commission, 2003), all companies in our data set are defined as large companies. However, as the company sizes range from 665 employees (subsidiary) for the smallest company to more than 540,000 employees for the largest company, a further subdivision by the number of employees is made to give a more accurate picture of the data set (see Table 3.5).

Range of company size	Distribution in %	Avg. # of CE programs
1) Companies with >500 and <1,000 employees	6%	1
2) Companies with >1,000 and <5,000 employees	11%	1
3) Companies with >5,000 and <10,000 employees	8%	1
4) Companies with >10,000 and <20,000 employees	19%	1.29
5) Companies with >20,000 and <50,000 employees	14%	1.20
6) Companies with >50,000 and <100,000 employees	17%	1.33
7) Companies with >100,000 employees	25%	2.33

Table 3.5: Distribution of company sizes in our data set

The data set shows a rather balanced distribution regarding the different company sizes, covering small, medium, and large companies. For the smaller companies up to 10,000 employees (1 – 3 in Table 3.5), we see that they operate only one CE program at a time. The companies larger than 10,000 employees (4 – 7 in Table 3.5) are operating on average 1.49 CE programs in our data set.

It must be stated that these numbers are based only on the cases we collected in our study. While for the smaller companies, it was possible to understand whether all CE programs were investigated, for the larger companies, the total number of CE programs operated by each company cannot be derived from the data. However, cross-checking with a recent consultant study (Lehnen et al., 2020) suggests that the number of CE programs operated by large companies is higher than the 1.49 from our data set.

D: Understanding the heterogeneity

In summary, this subsection has shown that the dataset exhibits high heterogeneity in terms of (A) industry classes, (B) ownership structure, and (C) company size. From a company perspective, the data set represents a cross-section of the industry landscape in the DACH region. This is important because there may be differences in the organizational design of CE programs depending on contextual factors such as industry class, ownership structure, or company size. Covering different manifestations of these factors increases the chance to better understand the heterogeneity of CE programs. Furthermore, it is expected that when the same configuration of CE programs can be found in a different organizational context, the validity of the identified CE program type is even stronger.

3.2.2 Data set from a program-perspective

In this subsection, background information on the CE programs themselves are provided. In the beginning, three superordinate CE program categories are introduced. These are used to present the information in a more detailed manner. The introduction of these categories is necessary for the following two reasons. First, there is nowadays a certain degree of uncertainty about the currently known CE programs (see Section 1.4 or Subsection 2.1.4). Second, the updated list of CE program types that will be developed to answer the first RQ will be presented in Chapter 5. Hence, to present the background information clearly, the three CE program categories appear to be the best way to do it. They are defined as follows.

1) *Internal CE programs*: fostering internal innovation initiatives by supporting intrapreneurial employees or by building new internal capabilities.
2) *External CE programs:* fostering innovation using different types of external means such as the collaboration with or investments in startups.
3) *Radical innovation units*: A collective term for organizational units that are combining multiple innovation activities or formats[8].

The dataset represents a cross-section of different CE programs that, to the best of my knowledge, reflects the full heterogeneity of organizational forms in CE. The 54 cases from our dataset fall into the three categories as follows.

[8] Examples for names used in practice are innovation lab, digital lab, digital factory, or innovation hub.

1) 23 cases are internal CE programs (from 21 different companies)
2) 20 cases are external CE programs (from 18 different companies)
3) 11 cases are radical innovation units (from 10 different companies)

Having presented the three CE program categories, next the background information will be provided. They will be distinguished by (A) the age, (B) the location, and (C) the size.

A: Age of the CE programs

For calculating the age of the CE programs, the end of October 2020 was chosen as the cut-off date. At that time, the youngest CE program that is still operating has an age of 1.83 years, whereas the oldest, still operating has an age of 24.75 years. The average age across the CE programs is 5.97 years (excluding the ones that were terminated).

Table 3.6 illustrates the age of the CE programs (in years) and compares two groups of (a) only CE programs that are still alive and (b) all CE programs, including the terminated ones.

CE program category	CE programs (still alive)			All CE programs		
	Age $_{MIN}$	Age $_{MAX}$	Age $_{AVG}$	Age $_{MIN}$	Age $_{MAX}$	Age $_{AVG}$
1) Internal CE programs	2.33	9.25	4.44	2.00	9.25	4.18
2) External CE programs	1.83	9.25	4.63	1.83	17.67	5.28
3) Radical innovation units	2.08	24.75	11.82	1.50	24.75	9.47
All CE programs	1.83	24.75	5.97	1.50	24.75	5.76

Table 3.6: Age distribution across the CE program categories

Age comparison of three classes

Below we compare the values of the three CE program categories from the "still alive" group. It can be observed that the internal and external CE programs have similar values for the minimum age, maximum age, and average age. In contrast, the values for the radical innovation units are more than twice as high, at 24.75 for the maximum age and 11.82 for the average age. A possible explanation for this is given on p.69.

For the group of "All CE programs", which also includes those that have been terminated, the general observation can be made that the overall age of CE programs is slightly decreasing. However, looking at each CE program category individually, it can be seen that the maximum age and average age of the external CE program category is increasing. In contrast, for the

internal CE programs and the radical innovation units, the average ages are decreasing. To better understand these different developments, the next step is to analyze the terminated CE programs in more detail.

In total, 8 out of 54 CE programs were terminated (or paused), which reflects 14.8% of all cases[9]. Four of them are internal CE programs, three are radical innovation units, and one is an external CE program. Comparing these numbers with the total cases of each CE program category shows the following rates of termination – 27.3% for radical innovation units, 17.4% for internal CE programs, and 5% for external CE programs. This circumstance will be discussed in more detail in Chapter 5 after the introduction of the different types of organizational designs to provide a more nuanced understanding regarding their differences.

In Table 3.7, the CE programs that have been terminated are compared with those that are still alive. The table shows that the overall average age of the group (a) [terminated CE programs] is lower than the ones of the group (b) [CE programs that are still alive]. However, there is one exception, namely the external CE programs in the group (a) which covers only one case with a comparably high age of 17.67 years. This exception also explains the observation from Table 3.6 that the age of internal CE programs and radical innovation units decreases when terminated cases are included, while the age of external CE programs increases.

CE program category	(a) CE programs (terminated)			(b) CE programs (still alive)		
	Age $_{MIN}$	**Age $_{MAX}$**	**Age $_{AVG}$**	**Age $_{MIN}$**	**Age $_{MAX}$**	**Age $_{AVG}$**
1) Internal CE programs	2.00	3.83	2.98	2.33	9.25	4.44
2) External CE programs	17.67	17.67	17.67	1.83	9.25	4.63
3) Radical innovation units	1.50	3.00	2.42	2.08	24.75	11.82
All CE programs	1.50	17.67	4.60	1.83	24.75	5.97

Table 3.7: Comparing age of CE programs that are terminated vs. still alive

[9] The number of terminated CE programs in our study does not reflect a general termination rate for CE programs. Due to our study design, we only interviewed CE programs that were operating. Of the 54 cases, eight were terminated by the end of October 2020. Examining the overall termination rate would require a different study design that captures both (a) CE programs that are still operating and (b) CE programs that have already terminated over a period of time.

Reasons for termination

Generally, the termination of a CE program is a strong indicator of the fact that it was not successful (see Burgelman & Välikangas, 2005). In order to better understand the differences in the age of the CE programs, a deeper look at the reasons for the termination of the eight cases has been made. For that investigation, we collected follow-up data (through desk research or follow-up messages with our interviewees).

This process revealed that five of the eight CE programs were terminated because they either (a) did not meet the expectations of the core organization or (b) the economic situation in the core business became worse. For the remaining three cases, the interviewees stated that the CE program was replaced by a larger one that was initiated due to the previous positive experiences. In this case, the activities of a CE program can be seen as rather successful since the company has understood its value.

Two reasons are more closely investigated and lead to a further distinction into the following two groups: (a) the ones that were terminated and perceived as being *unsuccessful* and (b) the ones that were terminated but replaced by a successor CE program, which was seen as *rather successful*. Moreover, the external CE program (age of 17.67 years) belongs to group b as it has been replaced by an even larger, company-wide program to intensify the collaboration with startups.

Figure 3.2 is showing these two groups and the respective average ages of the CE programs.

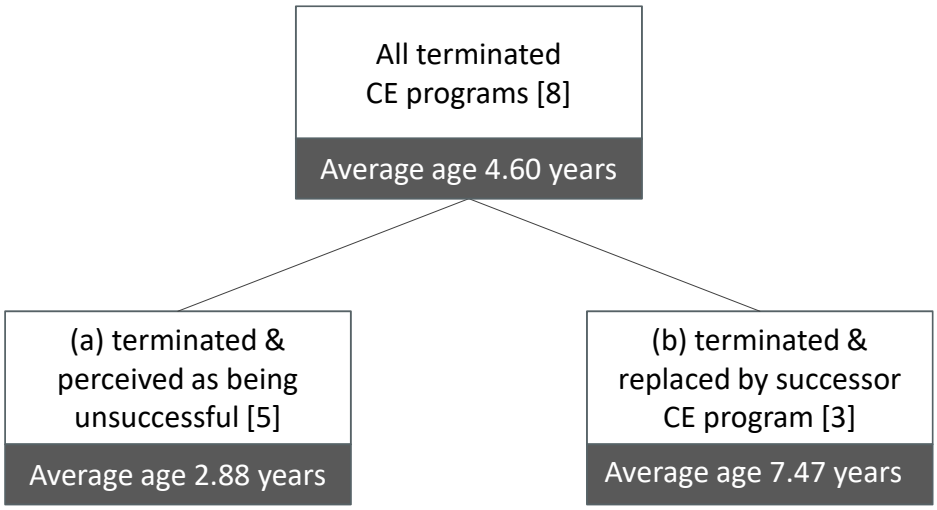

Figure 3.2: Different groups of terminated CE programs

Distinguishing the terminated CE programs into two groups shows that the ones that are terminated and perceived as being unsuccessful have an average age of only 2.88 years. This is significantly lower than the average age across all CE programs in our data set (5.76 years, see Table 3.6). The fact that terminated CE programs, which are perceived to be unsuccessful, have an average age of only 2.88 years supports the assumption that unsuccessful CE programs are terminated relatively soon after their initiation (Ma, 2020; Burgelman & Välikangas, 2005). Furthermore, it must be noted that not all CE programs that are closed are perceived as being unsuccessful, especially not when they are replaced by a successor CE program (see (b), with an average age of 7.47 years).

Subsection conclusion

This subsection has shown that the age of the CE programs shows a broad range from 1.83 for the youngest one and 24.75 years for the oldest still operating one (see Table 3.6). Furthermore, the data shows that between the ones that are terminated and the ones that are still operating, a strong difference in their age can be observed. In that respect, it is important to consider the reason for the termination of the CE program, as the cases in our data set indicate that not every CE program that is terminated is unsuccessful. However, for the ones that are being perceived as rather unsuccessful, the average age is much lower.

B: Location of the CE programs

Our study includes CE programs from companies headquartered in the DACH region and one from a German subsidiary of a large multinational corporation from Spain. A large proportion of the CE programs is located in Germany (39 cases); of them, one case belongs to a Swiss company. Only two cases are located in Switzerland and one in Austria. In addition, twelve cases are so-called innovation outposts located in Silicon Valley (USA), of which eleven are German companies, and the remaining one is Swiss.

In most cases, the CE programs were located close to the main office[10] of the respective company. The locations for the three CE program categories are as follows.

- *Internal CE programs*: 19 of the 23 internal CE programs were located near the main office. In the other four cases, the CE programs were located in innovation hotspots

[10] Some companies did have multiple main offices, e.g., when having several large business units, but only one headquarter. The term main office is used as a CE program is not necessarily located nearby the headquarter.

(two in Silicon Valley, two in Germany). In two cases, the internal CE program was located in multiple locations, including close to the main office.

- *External CE programs*: Twelve out of the 20 external CE programs were located near a main office in the DACH region. The remaining 8 CE programs did have their location in innovation hotspots (seven in Silicon Valley, one in Berlin), whereof four of them did operate multiple locations worldwide.
- *Radical innovation units*: Eight out of the eleven radical innovation units were located close to the main office. The three remaining ones were located in Silicon Valley.

Distinguishing the CE programs according to their location shows that there is a difference in age between the CE programs located in (a) the DACH region vs. the ones located in (b) the Silicon Valley. Table 3.8 illustrates the difference in age between the two geographical groups.

CE program category	(a) CE programs (DACH region)			(b) CE programs (Silicon Valley)		
	Age $_{MIN}$	**Age $_{MAX}$**	**Age $_{AVG}$**	**Age $_{MIN}$**	**Age $_{MAX}$**	**Age $_{AVG}$**
1) Internal CE programs	2.00	9.25	4.25	2.75	2.75	2.75
2) External CE programs	1.83	8.42	4.44	2.25	17.67	6.87
3) Radical innovation units	1.50	18.83	5.25	20.33	24.75	22.14
All CE programs	1.50	18.83	4.47	2.25	24.75	10.66

Table 3.8: Age distribution of CE Programs without innovation outposts

Subsection conclusion

The comparison of the two geographical groups (DACH region vs. Silicon Valley) shows that the CE programs located in Silicon Valley have a significantly higher age than those in the DACH region. The average age for all CE program categories is 10.66 years for those in Silicon Valley, more than double the age of those in the DACH region at 4.47 years. In addition, the following things can be stated about the two geographical groups.

1) The category of internal CE programs includes only one case located in Silicon Valley. This case is among those that were paused during the Corona crisis, which may explain why the average age of the CE program category is lower than those in DACH region.
2) Radical innovation units located in Silicon Valley show the highest values for minimum, maximum, and average age among all cases. With an average age of more than four

times that of the radical innovation units in the DACH region, they appear to be a major driver of the difference in average age between the two geographical groups.

C: Size of the CE programs

The third type of background information is the size of the CE programs. In general, it can be observed that the size of the CE programs shows a wide range, with the smallest program employing only one person, in contrast to the largest with 300 employees. The average size according to the three CE program categories is as follows.

- Internal CE programs = 10.93 employees on average
- External CE programs = 9.45 employees on average
- Radical innovation units = 75.82 employees on average

For explaining the differences between the average size of internal and external CE programs on the one hand and the radical innovation units on the other, a better understanding of the different organizational designs seems to be required. This understanding will be provided in Chapter 5 by introducing an updated list of the different types of organizational designs that have been identified in our study. Having defined these types, the size and the other background information will be described and discussed in a more detailed way.

Subsection conclusion

To conclude on the data set, the cases covered by the data set show a high degree of heterogeneity regarding the three types of background information (A) age, (B) location, and (C) size. This is supportive for understanding the heterogeneity of (1) the organizational designs and (2) the value creation of CE programs for the following two reasons.

Organizational design: A larger number of cases located in different geographic regions (DACH region and Silicon Valley) is helpful to understand the full range of different organizational designs, as they may be influenced by external contextual factors. In addition, looking at CE programs with different levels of maturity provides an opportunity to examine whether and how organizational designs change over time or if they are similar regardless of their age.

Value creation: With an average age of 5.97 years, the CE programs in our data set appear to be well suited for examining the value creation of CE programs, as it is expected that some outputs will require a longer time frame to become observable. Furthermore, the different locations of CE programs may be appropriate for revealing certain types of outputs that are

different or unique to the area in which they are located. For understanding value creation, the size of the CE program also seems to be an interesting element, as more available resources (i.e., number of employees) may influence the variety or intensity of outputs.

To conclude on Section 3.2, the data set seems well suited to answer the RQs of our study.

3.3 Data analysis

This section will focus on the data analysis of our study. A purely qualitative research method will be used to answer the RQs. Our data analysis will follow grounded theory principles, which is one of the most established and widely used approaches (see Bell et al., 2019). The grounded theory goes back to Glaser and Strauss and their work from 1967.

In Subsection 3.3.1, the three-step coding process will be described. Subsection 3.3.2 focuses on answering the RQ1 to RQ3. Subsection 3.3.3 is describing the actions that were undertaken to ensure the quality of our research.

3.3.1 Coding process following grounded theory principles

The grounded theory coding process follows a three-step process that is divided into (A) *open coding, (B) axial coding,* and *(C) selective coding*. Throughout the entire coding process, continuous comparison and theoretical embedding are important tools to improve the quality of the results. The three coding steps will focus on the aspects described below (see Bell et al., 2019; Walker & Myrick, 2006; Strauss & Corbin, 1994). After that, we describe the *application* of the three coding steps.

A: Open coding in theory

In the open coding step, the data (consisting of interview transcripts, field notes, and other secondary sources) will be analyzed line by line. The aim is to identify and conceptualize different concepts and phenomena that occur in the data. The result of this coding step will be a large number of codes that are into broader categories (higher-order concepts).

B: Axial coding in theory

The axial coding step focuses on identifying relationships and patterns between the codes and categories described in the open coding. Therefore, the codes are examined for causal conditions, context variables, interaction patterns, or consequences of actions or interactions.

C: Selective coding in theory

The selective coding step focuses on the core category, also called the core variable, which is the central topic of the research. The aim is to systematically link additional categories to the core and to identify and validate relationships between the categories. The result is a holistic understanding of the core variable. For this purpose, the data set is continuously expanded (theoretical sampling) until saturation is reached.

Having described the three general steps of the grounded theory, their *application* in the course of our study will be discussed below.

A: Applied open coding

The aim of the open coding step is to identify the different relevant themes for answering our RQs. This step resulted in the three themes of (a) organizational designs and design elements to describe CE programs, (b) value creation and outputs created by CE programs, and (c) contextual factors that seem to influence the work of CE programs. For the coding of the interviews and the additional data, we used the content analysis program NVIVO 10. The open coding resulted in 184 concepts (codes) that were identified, as counted by the author.

B: Applied axial coding

The aim of the axial coding step is to align and harmonize the codes that resulted from the open coding. The focus was set on the two themes of (a) organizational design elements and (b) the outputs of CE programs since they build the basis for answering RQ1 and RQ2. Due to the different starting points in terms of existing knowledge on the two themes, two different approaches were chosen in order to harmonize them.

- For the organizational designs of CE programs, several studies exist (see, e.g., Gutmann, 2018; Weiblen & Chesbrough, 2015). To ensure a good theoretical embedding, it is relevant to consider these existing results (Gioia et al., 2013). The concepts from open coding were aligned with design elements derived from literature, resulting in a set of 26 design elements.
- In contrast to the organizational design, for the value creation of CE programs, there are no previous studies that are comparing different types of value creation of CE programs. The concepts identified in the open coding step were aligned to create a set of outputs that are applicable across the different CE programs.

C: Applied selective coding

The aim of the selective coding is (a) coding all cases and (b) making a targeted enlargement of the data set. The coding coincides with the continuous comparison of the codes and the data. During the process, the codebooks for the organizational design elements and the outputs were continuously adjusted, which resulted in 26 codes for the design elements and 27 codes for the outputs of CE programs.

3.3.2 Answering the research questions

The coding process described above built the basis for answering the RQs of our study. The next paragraphs will describe how codes that were derived from the coding process are used to answer the RQs.

RQ1: What are the different types of CE programs?

The first RQ focuses on defining and distinguishing the different types of CE programs. The following steps were undertaken. First, the 26 codes (derived from the data and literature) were used to create a morphological box[11]. The box consists of (a) the design elements and (b) the element characteristics. For example, (ad a) innovation type is considered a design element, and (ad b) process innovation or product innovation are considered as two potential characteristics. For each of the 54 cases, such a morphological box was created. It serves to represent the specific organizational design of the case. Second, the cases were compared with each other. This resulted in the identification of cases that share a high degree of similar design elements, which represent one CE program type. Third, based on these cases and the design elements, for each CE program type, the following holds: (1) a definition was provided, and (2) design elements were highlighted that are suited to distinguish one type from another.

RQ2: What types of outputs are created by CE programs?

In contrast to the organizational designs, the codes for the value creation of CE programs do already represent the answer for the second RQ. The harmonized set of 27 outputs (codes) created through the coding of the interviews represents the different types of value creation through CE programs. To provide a clear answer for RQ2, the outputs were clustered into

[11] In general, morphological analysis is a structured approach for investigating a variety of possible solutions, which is suited for multidimensional, complex problems that are non-quantifiable Ritchey (2011b).

seven categories and described clearly. For further empirical embedding, they were complemented with quotes from the data.

RQ3: Can causal relationships between CE programs and their outputs be identified?
Having understood the different types of CE programs and outputs, they are now investigated for relationships between them. For the investigation of the relationships, different types of analysis, e.g., box analysis, coding comparison, and crosstables, were executed. The content management program NVIVO 10 was used to run different analyses automatically, allowing an easy comparison of different variables, which is helpful for identifying patterns and differences. This step resulted in the identification of different types of relationships between the CE program types and the outputs, which answers RQ3 by showing how they are related.

3.3.3 Ensuring the quality by five criteria

Ensuring the quality of research is a highly relevant topic. Meanwhile, it is acknowledged that the quality criteria for quantitative and qualitative studies are different (see, e.g., Bell et al., 2019; Frambach et al., 2013). In quantitative research, the focus is set on the criteria of *validity* and *reliability* (see Bryman et al., 2008), which, however, do not apply to qualitative studies (see Korstjens & Moser, 2018).

In general, quality criteria for qualitative research are frequently and controversially discussed (cf. Bell et al., 2019; Korstjens & Moser, 2018; Tracy, 2010; Bryman et al., 2008). In the course of this work, we follow the five quality criteria described below that are proposed by various authors (see Stenfors et al., 2020; Korstjens & Moser, 2018).

(A) *Credibility*: how trustable and believable are the findings to others?
(B) *Transferability*: is there a potential application of the findings in another field?
(C) *Dependability*: how consistent are the findings in relation to the context of the data?
(D) *Confirmability*: how clear are the findings embedded and derived from the data?
(E) *Reflexivity*: is the researcher aware of their impact on the research process?

For the five quality criteria, we did apply in a total of 18 actions which will be described subsequently. The respective actions are guided by recommendations from the literature on how to ensure quality in qualitative studies (see Stenfors et al., 2020; Korstjens & Moser, 2018; Treharne & Riggs, 2015; Frambach et al., 2013). A list of definitions for the eight terms highlighted in italics in this subsection will be provided in Appendix 3.

A: Ensuring credibility

For ensuring credibility, which focuses on the extent to which others can trust the results, the following six actions have been undertaken.

- Using multiple data sources, such as interviews, internal documents, company presentations, or press releases to achieve triangulation of data.
- Collecting follow-up data, either through interviews or messages, to resolve ambiguities and to collecting missing information in order to ensure a holistic and comprehensive understanding of each case.
- Discussing the findings with some of the interviewees to verify/falsify conclusions drawn from the interviews.
- Presenting the findings at practitioners' conferences to receive additional feedback and to strengthen the credibility among the practitioners to build more trustworthy relationships with the interviewees.
- Using recommendations from interviewees or academic colleagues to establish a good level of trustworthiness when conducting new interviews.
- Using quotations to underline findings and make them easier for the reader to understand.

B: Ensuring transferability

For ensuring transferability, which focuses on the degree to which the results can be transferred to other fields of application, the following three actions were undertaken.

- Documenting the research process and providing detailed descriptions of the data set, including various background information on the companies and the CE programs.
- Conducting a data set that is comparably large with 54 cases that are covering the full heterogeneity of CE programs.
- Having covered cases from the three countries of the DACH region and so-called innovation outposts located in the Silicon Valley. Different cultural and economic contexts of the cases strengthen the transferability to other countries that share similar characteristics.

C: Ensuring dependability

To ensure dependability, which focuses on the extent to which the results are consistent given the context in which they were generated, the following three actions were performed.

- Striving for theoretical saturation in two ways. First, by having covered all the different organizational designs that were known at that time (covering full heterogeneity). Second, by having collected multiple cases for each CE program type to better understand their differences and similarities.
- Iterating multiple times through the data and applying different perspectives when analyzing them. By frequent publications (at least one publication per year) the findings were additionally reviewed by other academics.
- Using semi-structured interviews with open-ended questions to ensure that a holistic understanding of the phenomenon was created. Preliminary interviewing training exercises and debriefing of interviews to ensure good quality interviews.

D: Ensuring confirmability

For ensuring confirmability, which focuses on how well the findings are embedded into the data (reducing the researchers' bias), the following four actions were undertaken.

- Using accurate transcriptions of the semi-structured interviews to capture nuances and slight differences in the various concepts. Furthermore, the use of a software tool for coding and analyzing the data.
- Discussing the results, new ideas, as well as any ambiguities in the data with colleagues from the research institute to ensure a more objective interpretation.
- Discussing results with a fellow researcher who has extensive experience in the field of corporate entrepreneurship to reduce own biases in data analysis and to strengthen the trustworthiness of the results.
- Conducting the first 15 interviews jointly with colleagues from the research institute to (a) increase the quality of the interviews through a post-interview feedback session and (b) reduce ambiguities in initial interpretations and the formation of first ideas.

E: Ensuring reflexivity

For ensuring reflexivity, which focuses on how aware researchers are regarding their own role within the study, the following two actions were undertaken.

- Visiting an annual, European-wide graduate school, to receive feedback from other Ph.D. students and experienced scientists and to reflect on my research activities.
- Collecting ideas, future research avenues, good practices, or open questions in a notebook that was used throughout the research process.

Subsection conclusion

The 18 actions described above were performed throughout the research process. The aim of them is to strengthen each of the five quality criteria (A to E) in order to ensure a high quality of the results of our study. Each of the five criteria is expected to contribute to a higher quality of the research (see, e.g., Stenfors et al., 2020).

However, there is no weighting for the five criteria, and consequently, it is rather difficult to evaluate the relative and exact contribution of each of the 18 actions that were undertaken. Nevertheless, the actual undertaking of these actions is recommended by previous studies (Korstjens & Moser, 2018), which will, in our opinion, contribute to a higher quality level and the *trustworthiness* of our results.