Chapter 5
Estimator Socialization in Design Thinking: The Dynamic Process of Learning How to Judge Creative Work

Julia P. A. von Thienen, Steven Ney, and Christoph Meinel

Abstract The assessment of ideas is a central activity in creative processes. Since teachers and coaches guide the learning of students, their assessment styles are particularly consequential. We report a longitudinal study, comparing the idea evaluation style of coaches before and after they are trained in the innovation paradigm design thinking. Initially, the coaches display a static idea assessment style. They attribute value primarily based on the requirement that ideas should be immediately effective, regardless of whether students are in the middle or in a late stage of their creative process. After being trained, the coaches have developed a dynamic, process-oriented evaluation style. They also assess ideas in line with design thinking values, with one exception. Contrary to design thinking teachings, the coaches do not come to value idea originality. The chapter closes with considerations how to facilitate the acceptance of original ideas.

5.1 Introduction

The assessment of ideas – finding ideas promising or futile – often decides which path a creative project takes. This happens for good or bad. When assessments go astray, fruitless ideas may be pursued with ample resources until they all too obviously reach a dead end, or ground-breaking ideas may be unnecessarily abandoned.

J. P. A. von Thienen (✉)
Digital Engineering Faculty, University of Potsdam, Potsdam, Germany
e-mail: Julia.vonThienen@hpi.de

S. Ney
HPI Academy, Potsdam, Germany
e-mail: Steven.Ney@hpi.de

C. Meinel
Hasso Plattner Institute at the University of Potsdam, Potsdam, Germany
e-mail: Christoph.Meinel@hpi.de

R. A. Beghetto, G. E. Corazza (eds.), *Dynamic Perspectives on Creativity*,
Creativity Theory and Action in Education 4,
https://doi.org/10.1007/978-3-319-99163-4_5

In creativity education the handling of ideas is taught both on explicit and implicit levels. What teachers say about the evaluation of creative ideas is one thing. How they handle student ideas is another. When teachers react to student ideas in approving or dismissive ways, they implicitly convey standards how to think and feel about ideas.

The assessment of creative ideas is unlike the assessment of student answers in a math test where teachers can "judge" what is right or wrong according to rather static standards. Corazza (2017) therefore suggests a change of language: In creativity education we don't *judge* the value of creative ideas, we *estimate* it. Yet, the question how teachers can learn to be good estimators of student ideas is still an open one, requiring further research.

> We must progress in the understanding of
>
> – how educators can be good estimators of creativity
> – how they can foster the development of estimation ability in their students, an essential part of their creative mindset
>
> (Corazza 2017, p. 22)

It has been repeatedly observed that educators who are untrained in creativity paradigms intuitively adopt idea assessment styles, which in fact counter the development of creative mindsets in students. For instance, at school teachers regularly perceive unexpected student ideas as disruptive (Beghetto 2007, 2010). Consequently, teachers are inclined to dismiss these ideas, hoping to ensure seamless on-task work in class. However, unexpected student ideas are potentially creative. To aid the development of creative mindsets, teachers should learn to explore these ideas more frequently (Beghetto 2013). Maybe creativity education for teachers could endow them with revised assessment standards, ideally yielding novel intuitions altogether. Unexpected student ideas would then be perceived as opportunities rather than threats by the teachers, if not always then at least more regularly.

Is this possible? Can estimators learn to perceive ideas in novel ways, such as to better facilitate creative processes and the development of creative mindsets?

In this chapter we report a longitudinal study, tracing how idea assessments of estimators change as they undergo training to become certified innovation facilitators. Study participants attend a one-semester *Certification Program for Design Thinking Coaches* at the Hasso Plattner Institute (HPI) of Potsdam University. Design thinking is an innovation paradigm taught at an increasing number of academic institutions world-wide, including Stanford, Potsdam and Cape Town University. At Potsdam, we presently teach 320 students each year in formats such as the "Design Thinking Weeks" (80 students), the one-semester long "Basic Track" (160 students) and the also one semester long "Advanced Track" (80 students).

In the *Manual for Design Thinking Coaches* (Ney 2016) participants of the *Certification Program* learn about their tasks as innovation facilitators. Preparing and hosting creative processes of design thinking teams are important objectives.

> Team coaches need to apply their judgement about when to adhere to and when to depart from the plan, when to leave the team be and when to intervene, as well as how to best help a team that has got itself stuck. […]

> [T]eam coaches need to forge *links* and create effective transitions from one design thinking phase to the next. Here, team coaches not only have to ensure that teams generate the outputs that enable them to address the tasks of the subsequent phase, they also need to support teams in creating the type of outputs that inspire and promote innovation.
> (Ney 2016, p. 9, emphasis in original)

Among other things, the *Certification Program for Design Thinking Coaches* attempts to school the estimation abilities of participants. After the training, coaches shall be expert estimators who reliably sense whether teams progress on trajectories with a high creative potential, or whether teams go astray and potentially need to be redirected by means of coaching interventions. Team coaches shall also estimate to what extent intermediate process outcomes – such as ideas – adequately promote the development of innovation, i.e. creative achievements, in subsequent process stages.

The basic objective of our longitudinal study is anticipated by Beghetto's insinuation: "As with all assessments, when it comes to assessing creativity, what you assess is essentially what you get" (2010, p. 453). This observation yields key questions we hope to clarify. What do people assess when they estimate the value of ideas? Do the assessment styles of estimators change in characteristic ways when people undergo training in a creativity paradigm?

We will first introduce idea characteristics, which are considered important in design thinking education (Sect. 5.2). We will then describe our longitudinal study, including the measurement approach that was developed to analyse assessment styles of estimators (Sect. 5.3). Then study outcomes will be reported (Sect. 5.4) and discussed (Sect. 5.5). As one study result indicates, the handling of idea originality may be especially difficult to learn and teach. This issue concerns creativity educators far beyond design thinking. We will close this chapter with considerations how to facilitate the acceptance of original ideas (Sect. 5.6).

5.2 Valuing Ideas Design Thinking Style

The design thinking community has gradually developed some consensus on how to think and feel about ideas. This consensus certainly does not seek to eliminate all judgement variance that people naturally produce when thinking about ideas. Different valuation perspectives are a tremendous resource to be cherished in creative communities. Instead, design thinking courses shall help trainees develop "antennas" for some aspects of ideas, or evaluative dimensions, to which they may have been insensitive before. Trainees can also expect to develop novel emotional preferences regarding ideas and to unlearn others.

The evaluative style that is characteristic of design thinking culture at present has been shaped over decades. It is informed by personal preferences of pioneering community members, philosophical positions, explicit argumentation and, increasingly often, by empirical studies. While the style is conveyed through implicit and explicit enculturation processes (Sects. 5.2.1, 5.2.2 and 5.2.3), it is not an arbitrary

culture. Instead, it is based on clear assumptions and is open to argument-driven revisions. In Sect. 5.2.3 we will discuss some questions where novel data could potentially impact the community's future way of assessing ideas.

We will review four dimensions, on which ideas can vary, to characterize design-thinking-typical ways of assessing ideas. While the discussion serves to clarify design thinking specific patterns of idea assessment, we hope to inspire readers beyond this particular creativity paradigm. The question what dimensions one invokes to estimate the potential of ideas is crucial for creative processes, whatever approach to creativity one pursues. Another important question is how the assessment standards are conveyed to novices, both explicitly and implicitly. Readers with a background other than design thinking are specifically invited to compare the assessment dimensions outlined below, and educational strategies used on their behalf, to the assessment dimensions and respective teaching approaches used in their own field.

5.2.1 Focus on User Needs

Already in the 1950s, when precursors to design thinking trainings began to develop at Stanford University, a humanistic philosophical orientation informed the emerging approach. John Arnold, who pioneered creativity education in the engineering department, discussed societal challenges as important starting points for creative endeavours (Arnold 1959/2016; von Thienen et al. 2017a). He prompted his students to tackle issues such as world hunger or traffic deaths. In Arnold's view, the task of engineers, designers, inventors and generally product developers is to identify bad conditions in the world and to bring about positive change by means of creative solutions. Based on this general philosophy, one of his predecessors in the department, Robert McKim, formulated a design theory based on human needs (McKim 1959/2016; von Thienen et al. forthcoming-a). According to McKim, the task of designers is to satisfy the physical, emotional and intellectual needs of mankind in morally and socially responsible ways. Present-day design thinking education continues these lines of thought. Creative projects evolve around user needs, which shall be gratified in socially responsible ways (Fig. 5.1).

Typically, in design thinking education students get to work in multi-disciplinary teams with three to six members. They work on real-life innovation challenges often provided by external project partners. The task can be, for instance, to create a better airport experience, which may be the wish of an airport operating company. At first the students shall *understand* the airport domain by conducting research about it. In the *observe* phase the students interview travellers, security guards, check-in personnel or even pizza sellers at airports and make careful behaviour observations; they can also build on personal airport experiences. In the *point of view* phase the team specifies key insights and decides who their user shall be. They may decide to create a better airport experience for travellers, in which case security guards, pizza sellers etc. will not be the addressees of their project henceforth. The team then

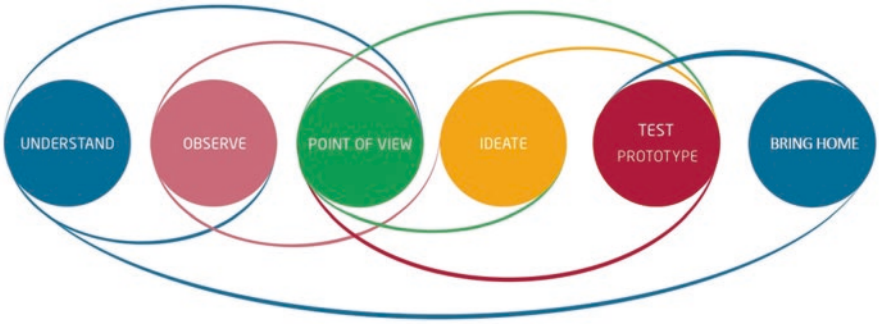

Fig. 5.1 The design thinking process

creates an imaginary customer to facilitate ideation, a so-called "persona". This could be Mrs. Wiggs, 62 years old, who flies regularly but finds it increasingly stressful to wait in long lines where she misses opportunities to sit down. In the *ideation* phase the team begins to think up solution ideas. They seek solutions for Mrs. Wiggs, considering as many and as diverse ideas as they possibly can. Afterwards, one or more ideas are selected. The team may decide to focus on the idea of a trolley to improve the experience of airport security checks: The trolley allows customers to separate fluids, technology, shoes etc. without hurry while lining up for the security check and also provides convenient seating. To test the idea, the team builds a prototype and tests it with persons who resemble the persona. User experiences shall be the ultimate criterion for the team to learn and advance their project. Team members should not defend a solution if users had unpleasant test experiences. In that case, the team should instead learn from the test and iterate their solution, returning to previous process phases if necessary. Finally, in the *bring home* phase a successful prototype is further advanced and potential implementation barriers are tackled.

A strong focus on user needs is taught explicitly basically in every session of design thinking education. The whole creative process, and methods used along the way, all support this purpose. In addition, design thinking novices can learn implicitly about this assessment dimension. Feedback from teachers or more experienced team members usually conveys how addressing user needs is of highest priority.

When a team ideates with a focus on user needs they seek solutions that specifically aid their intended user. In the airport scenario, the team seeks solutions for Mrs. Wiggs – not solutions for check-in personal or security guards at the airport and certainly not solutions, which team members may find "cool" for personal reasons while the approach would disregard the specific situation of Mrs. Wiggs.

Such a strong and consistent focus on user needs has a number of advantages in the creative process.

- **Solution effectiveness:** A consistent focus on user needs ensures that the intended audiences (users) find the outcome of the creative process valuable.

- **Intrinsic motivation:** Students feel their projects warrant much personal effort. They see themselves working towards truly desirable ends (Brown and Katz 2009; Kelley and Kelley 2013) and find the challenges personally meaningful (McKim 1972). Thus, students build up intrinsic motivation and drive (von Thienen et al. 2016, 2017a); they work energetically on their task even when facing difficulties along the way.
- **Teamwork:** The concern for user needs helps teams establish a joint focus and pursue a joint goal. The process is not about "what I want" versus "what you want". Instead, all team members decide together for whom they seek a solution and then the group forges ahead jointly to deliver the best possible outcome for their selected user.
- **Social connectedness, self-efficacy and agency:** The focus on user needs entails empathy and collaboration. Design thinkers experience social connectedness to team mates and users, which is an important resource in creative processes. As Cojuharenco et al. (2016) have demonstrated, social connectedness promotes self-efficacy (the belief of a person that she can make a difference) as well as agency (her taking of action even if positive effects are not immediately visible). This is especially important in creative projects. After all, creative achievements may materialize in the end, but people need to take action with no guarantees of success (Corazza 2016a).
- **Testing opportunities:** The focus on user needs introduces excellent opportunities for creative teams to test and learn. Users should embrace novel solutions like "revelations": Their gnawing, unmet needs would finally be satisfied. If test users do not celebrate a presented prototype, the creative team has something important to learn.

Despite of these advantages, the focus on user needs is not completely uncontroversial in the community. Occasionally people voice their preference that there should be more freedom to pursue personal visions, interests and intuitions. It should be possible to work more in the way artists proceed without having to focus on someone else's user needs. While such art-inspired process models might play a greater role in the future, as they did some decades ago, at present the user-focus is a clear and quite characteristic learning objective in design thinking education.

5.2.2 Balancing Team Interests

Working in teams is everyday-business in many product development units of the industry. Rarely does one individual have all the necessary skills to make inventions for a company alone, where software, hardware, usability and marketing expertise may be required. Such an industry inspired team-based work approach was also adopted early on in Stanford's creativity education for engineers, which is a major root of present-day design thinking. However, initially the team based approach was

Fig. 5.2 Explicit mottos convey design thinking values. Here, the motto "build on the ideas of others" is printed in large letters on the wall. (Photo by Toni Mattis)

just one training mode amongst others. In the late 1990s the emphasis on teamwork was strongly increased (Carleton and Leifer 2009). At present there is no teaching of individuals any more. Design thinking is completely team-based. Throughout their training students work in teams. Analogously, teachers teach in teams (Kelley and Kelley 2013; Roth 2015).

The value of collaborative invention is conveyed in multiple ways in design thinking communities. Explicitly, mottos such as "build on the ideas of others" prompt teams to ideate jointly (see Fig. 5.2).

Implicitly, students learn to think and act collaboratively both by what they witness and by what they don't witness. Experienced design thinkers act as role-models who live and teach in a collaborative spirit. In addition, it is also noteworthy what happens seldom in design thinking environments. Historically, theories of creativity often evolved around figures of "individual creative geniuses" (Beaney 2005; Gaut 2010). However, design thinking communities rarely narrate "hero-stories" of individual inventors. Rather, the community tells stories about both the creative achievements and struggles of creative teams (see, for instance, the collection of design thinking case stories on thisisdesignthinking.net, Hasso Plattner Institute for Digital Engineering 2017).

In creative teamwork, the question how to handle varying interests is a regular issue to come up. For instance, what if some of the team members want to build a technical solution whereas others favour a social solution? In this case, should the team maybe seek a social and technical solution? After all, such a combined approach could help keep all team members engaged in the process.

Coaching novices sometimes interpret the literature on teamwork and social competence as promoting the integration of all team member interests as a

high-ranking requirement. Yet, experienced design thinkers do not place questions regarding team member interests centre-stage in the ideation phase. When teams seek solutions, they shall "saturate the solution space", considering as many different approaches as they possibly can. Limiting the solution space to approaches that satisfy all team member interests would seem counterproductive (e.g., considering solutions only that are social and technical). Furthermore, solutions shall be tailored to user needs, not to team member interests. Finally, teams learn to "bias toward doing and making over thinking and meeting" (d.school 2010, p. iii). Teams can trust that subsequent user tests will clarify the potential of ideas. Arguments about the issue featuring personal preferences seem highly unnecessary.

Thus, ideas that balance team interests are at present only slightly preferred in design thinking communities. While social skills are generally accorded great importance, the evaluation of solution ideas should not be overly limited by the need to accommodate diverse private interests of team members.

5.2.3 Originality and Effectiveness

Creative achievements are commonly defined as outcomes of creative processes that are *original and effective,* or, in another parlance, *novel and valuable* (Gaut 2010; Runco and Jaeger 2012; Corazza 2016a). These notions will be discussed in more detail below as they had to be operationalized in our study (see Sect. 5.3.1). For now we trust that a considerable consensus has been achieved as to what the terms mean.

In virtue of being an innovation paradigm, design thinking carefully attends to the originality and effectiveness of solution ideas. Thus, design thinking shares these two evaluative dimensions with many other approaches to creativity and innovation, which helps to draw from a rich corpus of theories and research results, and to jointly advance knowledge in the field.

One resource to draw from is a model of thought trajectories provided by Corazza (2015, 2016b) as part of the *Dynamic Creativity Framework.* The model depicts typical ideation moves in creative processes, illuminating likely time-dynamics of innovation projects. It also provides explanations why the objective of creative projects, to arrive at original and effective outcomes, is often difficult to achieve. Notably, we live in sophisticated cultures where a lot of ideas have been contemplated before. Moving beyond those ideas that someone else has already thought up is therefore a non-trivial task that usually requires time. Persons who enter a creative process can expect to tap a number of non-novel options first. Then, once ideation trajectories enter the realm of novel ideas, effectiveness becomes particularly hard to achieve. Most novel ideas may seem funny, foolish, mad etc. Working out a creative breakthrough – an idea no one else has had before, which in addition proves utterly effective – needs to be acknowledged as a great feat. If creative breakthroughs ever materialize in a project, it is typically after a long creative process where the originator had to persist in the face of numerous inconclusive outcomes, i.e. ideas that did not seem sufficiently original and effective.

Corazza's trajectory model hints at interdependencies between originality and effectiveness. The two dimensions are likely to be non-orthogonal in most natural settings. Effectiveness seems easily achievable with conventional solution ideas. One simply replicates already existing approaches, which should be at least somewhat effective. When, by contrast, untried possibilities are explored, effectiveness is often rather difficult to reach.

Empirical evidence seems to support this view. In a study at the HPI Potsdam (von Thienen et al. 2011), 40 participants worked on a real-life innovation challenge over 1 week. Half of the participants had participated in design thinking courses. The other participants were interested in creative work but had not yet been enrolled in design thinking classes. Design thinking students developed significantly more original ideas compared to untrained students. However, when the effectiveness of outcomes was estimated by domain experts, a negative correlation of −0.55 obtained between idea originality and idea effectiveness, which was statistically significant at a level of $p \leq .01$. Considering only the most original ideas, i.e. those of design thinking students, this negative correlation even amounted to −.70. This data supports the view that originality and effectiveness are non-orthogonal dimensions in many natural settings. The higher the level of originality, the more difficult it becomes to achieve effectiveness.

Similar results were found by Agnoli et al. (under review) who studied patterns in the advertisement domain. They found that, generally, greater originality entailed less recognized creative achievements at work. However, work experience seemed to mediate between originality and creative achievement, so that experienced advertisers in effect benefit from more original ideas. Again, the picture emerges of a *prima facie* negative relationship between originality and effectiveness. However, the difficult task of achieving effectiveness with original ideas seems to be manageable by those persons who can build on a great amount of work experience.

There can be a number of reasons why original ideas face effectiveness-hurdles (von Thienen et al. 2017b). Original ideas may be more difficult to communicate: Audiences may struggle to understand novel concepts, including their value. Furthermore, some – or many – audiences seem to reject original ideas, regardless of the idea content. Blair and Mumford (2007) demonstrate this phenomenon in a study with two samples of undergraduate students. A first group of students generates ideas, then another group of students evaluates the suggestions. In addition, four independent expert raters characterize all ideas on a number of dimensions, such as idea originality, adherence to social norms or expected implementation effort. Notably, the study participants display a strong preference for unoriginal ideas. The authors even speak of an "undeniable disdain for […] original ideas" (p. 215). They explain the phenomenon in virtue of forecasting difficulties. Highly original ideas make it difficult for audiences to predict the ensuing effects. As the ideas are unprecedented, it is unclear whether hurdles will emerge in the implementation process and social effects are hard to foresee; they could be positive or negative. Thus, the authors describe a contrast between original ideas on the one hand and ideas yielding clear social benefits as well as ideas yielding predictable positive

short-term effects on the other. Audiences are said to prefer socially beneficial and immediately effective ideas over original ones.

In addition to cognitive difficulties of anticipating the effects of original solutions, straightforward emotional reactions to novelty can also play a role. In animal innovation research, *neophobia* (being afraid of novel things) as opposed to *neophelia* (being attracted by the novel) is discussed as an important parameter (Greenberg 2003; Kaufman and Kaufman 2014). In human creativity studies, Barron (1955) made related observations. He assessed 100 men and found stable patterns in their orientation towards originality. Some men displayed an emotional preference for simple and conventional stimuli; these men also gave conventional rather than original answers in a number of tests. Other men who preferred more complex, unconventional stimuli also produced more original answers in the test situations. Against this background it is easy to see how audiences of people who emotionally prefer the conventional, who display neophobia rather than neophilia, can make it very difficult for any original idea to achieve full effectiveness. These audiences can function as gatekeepers who reject original ideas and counter their implementation.

Finally, original ideas face a practical, marked disadvantage compared to established ideas when it comes to effectiveness assessments. Established ideas benefit from longer periods of refinement and people are more practiced in their application. To illustrate difficulties novel approaches have to overcome, the high jump may serve as an example. Some decades ago high jumpers used the so-called scissors technique. There is a notable span in how high people can jump with this technique, depending on how practiced they are and how refined their technique is. Then, at some point a person (be that Fosbury or yet someone else) decided to try jumping backwards, which is nowadays known as the flop technique. Today we believe that it is possible to jump higher with the flop technique than with scissors. However, upon first trying it out with an unrefined technique and no practice in its application, surely the person who tried jumping backwards for the first time would not jump as high as professionals using a refined scissors technique. Thus, novel approaches may seem less effective than established approaches in first tests, even if their potential is actually greater (Fig. 5.3).

All in all, the degree of immediately realizable effectiveness is often a bad predictor for potential long-term effectiveness. A novel approach may need refinement to work out; people may need practice to use it effectively. Furthermore, a novel idea may be difficult to understand for others, so that better communication strategies need to be developed first. Also, audiences may need to be exposed to a novel idea for some time to gradually become more familiar with it and thus feel more comfortable about it. This presumed loose linkage between short term and long term effectiveness is of course highly relevant for creativity education.

In the history of design thinking, the belief was adopted early on that ideals of originality and effectiveness should play varying roles over time in creative processes. In the ideation phase, originality is considered a “must-have”, whereas immediate effectiveness is only “nice-to-have”. Key beliefs in this regard were adopted from the brainstorming pioneer Alex Osborn. John Arnold was personally

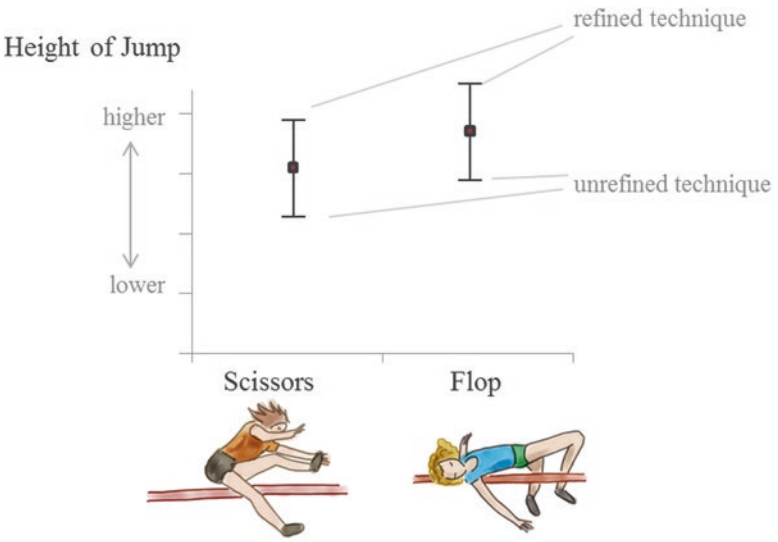

Fig. 5.3 Even if the flop technique has a greater potential effectiveness than the scissors technique in the high jump, it may seem less effective in first tests. Initially, an unrefined novel approach competes with a highly refined old approach. (Figure adapted from von Thienen et al. 2017b)

well-familiar with Osborn's approach, which he discussed in great detail and also critically (Arnold 1959/2016). Nonetheless, regarding originality and effectiveness Osborn's teachings have been maintained mostly unchanged in the design thinking community up to the present. "Osborn claims that it is easier to tame down than to think up" (Arnold 1959/2016, p. 106). Consequently, coaching instructions favour original ideas in the ideation phase. Arnold recapitulates what Osborn would say: "Remember now, men, we want as many ideas as possible – the wilder the better, and remember, no evaluation" (p. 105). Here, the instruction to avoid evaluation is given because evaluative thinking is taken to hinder the generation of multiple and original ideas. Up to the present, design thinking trainees still learn to "encourage wild ideas" and to "defer judgment" during ideation (d.school 2010, p. 28).

While these instructions are deeply engrained in design thinking practices, upon a closer look they actually seem contradictory. On the one hand people shall seek wild ideas, on the other hand they shall not evaluate. However, to seek wild ideas people must evaluate ideas on the dimension of originality, preferring original over unoriginal ideas. The instruction to refrain from evaluation obviously intends only one particular evaluative dimension, namely immediate effectiveness. This dimension shall play no important role in the ideation phase.

Given that design thinking is an innovation paradigm its projects must arrive at original and effective solutions eventually during the project term. This is rendered possible by a dynamically refined assessment strategy. Design thinkers do not aspire to originality and effectiveness uniformly throughout the process, which would mean to statically maintain one assessment style all throughout the project. Instead, assessment strategies undergo a fine-tuned regulation based on the following logic.

(0) Ideation is the process of thinking up ideas.

(I) Innovation requires original ideas.
(II) After ideation, the originality of ideas does not increase.
(III) Therefore, ideation must target original ideas to promote innovation.

(IV) Innovation requires effective ideas.
(V) When ideation targets immediately effective ideas, it tends to produce unoriginal ideas.
(III) Ideation must target original ideas to promote innovation.
(VI) Ideation need not target immediately effective ideas to promote innovation; the effectiveness of ideas can easily be increased through iterative prototype tests after ideation.
(VII) Therefore, ideation shall not target immediately effective ideas.

These arguments include empirically testable beliefs and more related research is likely to emerge. Notably, the development of idea effectiveness has already been investigated (claim VI).

In the design thinking process model (see Fig. 5.1), ideation is followed by the phase of testing prototypes. In practice this entails fast and highly iterative work routines. Prototypes are built quickly, tested and revised based on trial experiences. Indeed, this approach appears to be a reliable strategy to increase the effectiveness of basically any ideation outcome, as an experiment by Dow and Klemmer (2011) suggests. In their study 28 participants were asked to build vessels from everyday materials in 25 min to protect a raw egg from crushing that would be dropped from increasing heights. Task performance was measured in terms of the highest height at which the egg of each participant survived the fall. In the experimental condition, participants received a full carton of eggs and were prompted to test their vessel prototypes roughly every 5 min. Participants in the control condition received only one egg altogether. In the end, the average drop height that eggs survived was almost twice as high in the experimental condition compared to the control condition. Notably, solutions in the experimental condition did not excel because the participants switched their general approach based on test experiences. Rather, almost all participants stuck to their first ideation outcome (e.g., choosing a parachute, capsule or pillow design) and then only improved details of their approach. Thus, irrespective of the initial ideation outcome, quickly iterating prototype tests seem an excellent means to achieve high solution effectiveness. This supports the view that ideation need not target immediately effective ideas, since ideation outcomes can easily be rendered more effective in subsequent process phases.

At present, design thinking education conveys straightforward messages about originality and effectiveness as important idea dimensions. Explicitly students are prompted to produce original outcomes in all process phases up to ideation. In the understand and observe phases they shall gain new insights about user needs, which go beyond the explicit knowledge of humanity at the project outset. The teams shall

discern truly existing need-patterns in the project domain that no one could see or explicitly describe at the project outset – neither the team members, nor domain experts, nor the users. Then design thinking teams are expected to specify their point of view in a single sentence, describing what the user needs based on a key observation insight (d.school 2010, p. 21). Design thinking students often get to hear that their point-of-view-statement should be "a sentence no one else has ever thought before". Afterwards, in the ideation phase, teams are told to head for "wild" ideas, which means that they shall try to produce original ideas, not shying away from suggestions that may sound crazy or unrealistic when taken literally.

On behalf of effectiveness, students learn both explicitly and implicitly that this is a matter of addressing basic user needs. Consistently, in terms of explicit and implicit messages, students also learn to not choose ideas based on expectations of immediate effectiveness in the ideation stage. Wild-sounding ideas can be rendered more effective by means of testing and iterating prototypes in later project stages.

5.2.4 Study Hypotheses: From Static to Dynamic Assessment Styles

Definitions of creativity in terms of original and effective outcomes court a static view on creativity (cf. Corazza 2016a). Yet, creative processes can be better understood and supported from a dynamic perspective, as for most of the time original and effective ideas are sought, but they have not yet materialized.

When a static assessment style is endorsed, originality and effectiveness matter invariably, regardless of the stage a creative project is in. By contrast, dynamic assessment styles allow a careful orchestration of idea evaluations over time. Estimators can emphasize or de-emphasize assessment dimensions depending on the process stage.

Both static and dynamic assessment styles may exist in different versions. Design thinking is typically used in product-developing contexts. Such an environment is likely different from, e.g., the realm of art where by default originality likely plays a greater role.

In product-developing contexts "breakthrough innovation" may be an ideal, but most importantly novel outcomes shall be effective. In everyday business, radical change is rare. Incremental change is more common. When new products are developed, originality is often not even pursued as a self-standing goal, only increased effectiveness counts as progress.

We expect participants in our longitudinal study to initially display a static evaluation style, adhering to the typical values in product-developing contexts. Thus, even in the middle of the creative process – when considering ideation outcomes – we expect participants to react as though they were evaluating final products, assessing primarily the immediate effectiveness of ideas. After the training, we expect

participants to have developed a dynamic assessment style attuned to design thinking practices and values.

5.3 When Estimators Undergo Training: A Longitudinal Study of Idea Assessment Styles

To illuminate how people evaluate ideas in creative processes and how their evaluative style changes over time we developed *Idea Assessment Probes* (Sect. 5.3.1). After introducing this assessment methodology, we will describe the sample of study participants (Sect. 5.3.2) and the study procedure (Sect. 5.3.3).

5.3.1 Measuring Evaluation Styles with Idea Assessment Probes

The assessment methodology for this longitudinal study was developed in light of two constraints. First, the method should allow non-conflated, quantitative analyses regarding idea characteristics that impact value-judgements of single estimators. As discussed above (Sect. 5.2.3), in natural settings evaluative dimensions are often conflated. In particular, ideas that are more original typically show less immediate effectiveness. As Blair and Mumford (2007) would emphasise, the social benefits of original solutions also tend to be less clear. Forecasts are more difficult; both positive and negative consequences could obtain. Against this background, our assessment approach should render it possible to calculate the impact of single idea characteristics on value-judgements of estimators in non-conflated ways. In particular, it should be possible to assess the impact of idea originality on judgements of idea value without spurious correlations; high idea originality should not go along with reduced immediate effectiveness or reduced social benefits / less gratification of user needs. As a second requirement, assessments should be time-efficient. Study participants should not have to invest more than 10–12 min to fill out the questionnaire.

In light of these constraints, we developed *Idea Assessment Probes (IAPs)* as a measurement approach. Questionnaire items refer to ideas, which vary systematically on the dimensions of interest in dichotomizing ways. E.g., the ideas to be evaluated by the participants are either clearly original or clearly unoriginal. Since our study concerns four idea dimensions, we had to generate $4^2 = 16$ ideas (see Fig. 5.4) to cover all possible combinations.

We created two *IAPs* altogether, one idea set for the pre-test and another idea set for the post-test. Each questionnaire consists of an instruction sheet with a short scenario description and 16 idea cards. The scenario suggests a situation right after ideation, thus in the middle rather than at the end of the creative process. This should

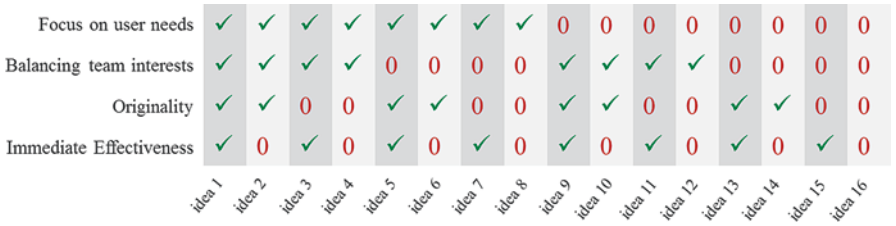

Fig. 5.4 Ideas in the questionnaire vary systematically on the dimensions (i) focus on user needs, (ii) balancing team interests, (iii) originality and (iv) immediate effectiveness

be a highly time-efficient method to distinguish between static and dynamic assessment styles of study participants.

To assess how strongly the value judgements of study participants are informed by design thinking practices and values, our scenario describes a persona (the indented user) and mentions diverging team interests. For instance, in the pre-test the following scenario is used.

You work as a design thinking coach. Your team has come up with persona Fritz Freundlich. Fritz is a 30-year old passionate farmer, who often suffers from back pain after harvesting asparagus, and who feels a bit lonely in his job. Your team has different preferences as to how the challenge should be tackled. Some team members want to concentrate on psychological wellbeing, while other team members want to concentrate on physical wellbeing.

To generate idea probes for the study participants to evaluate, we brainstormed about three times as many items as we finally included in the questionnaire. Two design thinking experts coded each idea on the four dimensions of interest. Only those ideas were considered for the questionnaire where the coding displayed perfect inter-rater-agreement. Among all remaining items, ideas were selected such as to avoid duplications of similar solution approaches, favouring instead a variety of differing solution ideas.

We used the following specifications to clarify the meaning of terms.

- *Focus on user needs*: The idea attempts to satisfy one or more persona needs; it does not conflict with persona interests.
- *Balancing team interests*: The idea addresses multiple needs; it accommodates the varying interests of different team members.
- *Originality*: The idea is novel. No product or service like this currently exists, nor is the idea familiar from fiction novels or movies.
- *Immediate Effectiveness*: The approach seems realizable without major hurdles. It is very likely to produce the intended effect – specified in brackets on each idea card – rather quickly.

One sample idea (# 4) is the following:

The farmer publishes an ad in a magazine to find a marriage partner. She can help him at work (so he has less back pain and is also less lonely).

This idea focuses on user needs because it attempts to satisfy the specific needs of Fritz Freundlich. The approach balances team interests because it pursues both the user's physical and emotional well-being, which is what different team members want to do. The solution is unoriginal because publishing ads to seek marriage partners is not a novel thing to do. Furthermore, the plan of relieving Fritz from back pain and loneliness by finding a spouse via ads is unlikely to work out quickly and smoothly, so the idea is not immediately effective.

On each card the test-takers read about one idea and shall answer two questions.

- *How do you find this idea?* Answers can be provided on a five-point scale ranging from "terrible" to "excellent". Later answers are coded on a scale from −2 (terrible) to +2 (excellent), so that positive values indicate approval while negative values indicate disapproval.
- *If the team wants to pursue this idea as their only prototype, do you want to intervene?* Here, answers can be provided on a five-point scale ranging from "not at all" to "absolutely". Answers are then coded on a scale from −2 (not at all) to +2 (absolutely), so that positive values indicate the coaches' wish to intervene. Negative values indicate that the team shall proceed without coaching intervention.

Notably, this assessment strategy splits the creativity criterion of *effectiveness* in different aspects:

- *Subjective value* is operationalized via the question "How do you find this idea?" to be answered individually by each study participant.
- *Immediate effectiveness* is operationalized as the consensus expectancy of design thinking experts that a solution approach will produce the intended effect rather quickly and smoothly.
- *Long-term effectiveness* of solution ideas could allude to the realisability of the intended effect by means of a prolonged process, in which hurdles may have to be overcome. Yet, this aspect is not assessed in our present study.
- Furthermore, ideas may have *serendipitous value:* They can bring about additional benefits beyond the originally indented effects. This aspect is also not assessed in our study.

Figure 5.5 shows the complete item set of one test-taker.

We created questionnaire versions in German and English, which are the two most frequently spoken languages in the *Certification Program for Coaches* at the HPI Potsdam. Idea cards were cut out and hand-mixed to ensure that test-takers view them in a random order. The complete *Idea Assessment Probes* from our pre- and post-test as well as a discussion of their psychometric properties are provided by von Thienen et al. (forthcoming-b).

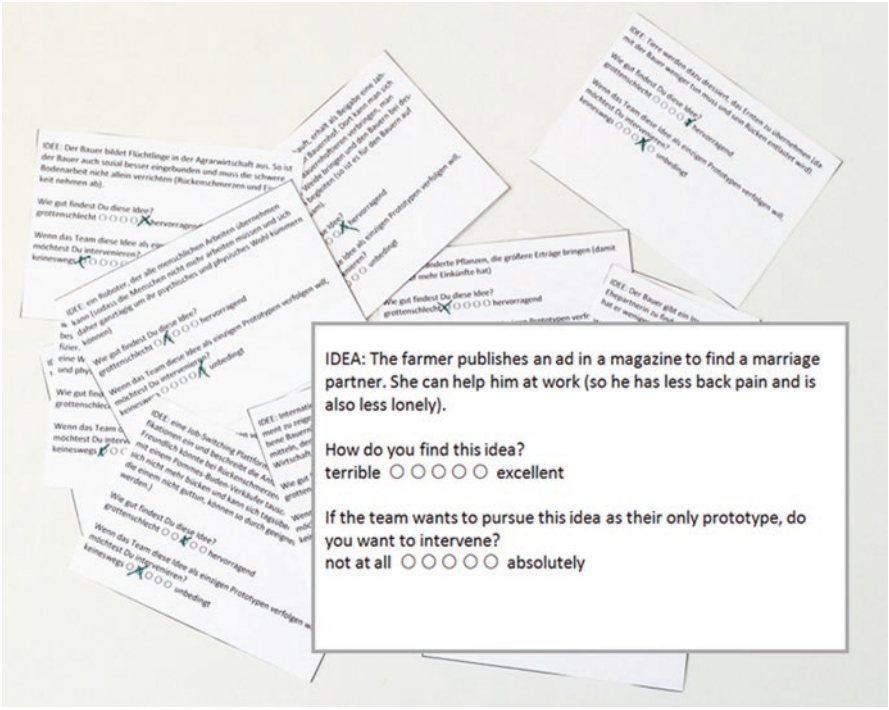

Fig. 5.5 The *Idea Assessment Probes* come in the form of 16 idea cards. One sample card is zoomed in

5.3.2 Study Participants

Each semester 25 coaches can participate in the *Certification Program for Coaches* at the HPI Potsdam. At our first assessment session 24 persons were present and thus included in the study, 16 males, 8 females. Their age ranged from 28 to 51, averaging on 39. Most participants were rather inexperienced with the design thinking approach and even fewer participants had prior coaching experiences (see Fig. 5.6).

5.3.3 Assessment Procedure

Design thinking sessions begin with warm-up exercises that specifically serve the purpose of advancing specific moods and cognitive styles, which are considered favourable for subsequent design thinking objectives. Therefore, care had to be taken to avoid conflated measurements in the pre-test. When the semester's *Certification Program* started, the program head briefly welcomed all participants

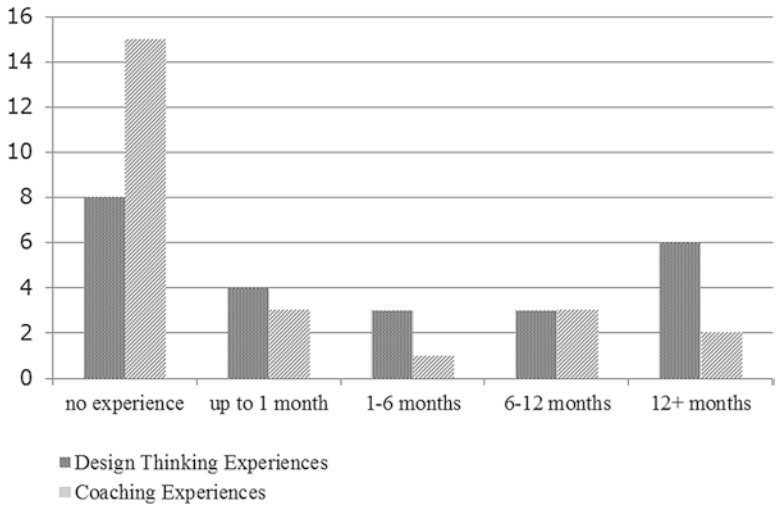

Fig. 5.6 Upon entering the *Certification Program,* the study participants are rather inexperienced in design thinking and coaching

for about 2 min, then immediately invited the attendees to take part in our study and questionnaires were given out. Participants could choose between German and English versions. A number of participants expressed their convenience with either language; they simply picked the physically closest version. In the pre-test, 14 participants filled out German questionnaires, 10 worked on English versions. All participants filled out the questionnaires completely, without producing invalid or missing answers. The assessment procedure took about 12 min.

The post-test was carried out in the final week of the *Certification Program*, following an analogous routine. Participants were welcomed in the morning by the head of the program for about 2 min. Then questionnaires were offered in two languages. In the post-test, 17 participants filled out German questionnaires and 7 attendees worked on English versions. Again, the participants provided valid answers on all questionnaire items.

To avoid redundancy we will not report the N (number of cases) in statistical analyses below, since it is always the same N = 24, the number of persons included in the study.

5.4 Results

Each questionnaire comprises 32 items: 16 ideas are provided and two questions asked about each of them.

In the pre- and post-training assessments different ideas were presented for evaluation, resulting in 64 items altogether being processed by the participants in the course of the whole study. Since the participants could choose between English and

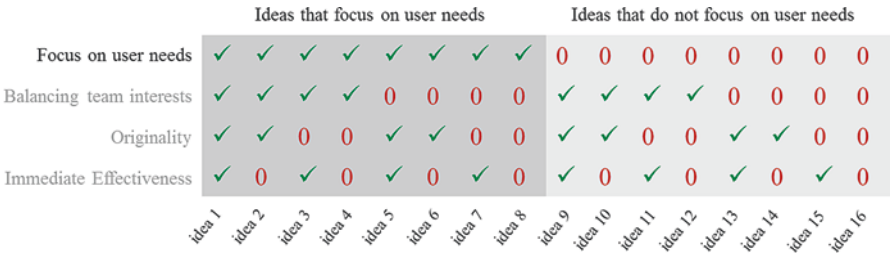

Fig. 5.7 On the four dimension of interest, group means are compared that draw on eight items each. To assess whether the idea characteristic "focus on user needs" impacts the coaches' responses, mean response values on items 1–8 are compared to mean response values on items 9–16

German questionnaire versions, we first assessed whether all answers would be comparable or whether the chosen language would impact participant responses. No statistically significant difference of answer-values was found on any of the 64 questionnaire items, which indicates a sound degree of test reliability.

We then pursued two major lines of data analysis. First, we calculated paired sample t-tests to see whether the four idea characteristics of interest (focus on user needs, balancing team interests, originality and immediate effectiveness) would impact the perception of idea value and the coaches' inclinations to intervene. For each t-test, the questionnaire's 16 idea items were split up in two groups to calculate means (see Fig. 5.7).

Furthermore, we calculated linear regression models, predicting the coaches' perception of idea value based on the four idea characteristics.

5.4.1 Focus on User Needs

At the beginning of the coaches training, the participants do not attend to the factor "focus on user needs". They value ideas that focus on user needs as much as they value other ideas (see Table 5.1, results at t_0).

By contrast, after the training, ideas that focus on user needs are clearly preferred (see Table 5.1, results at t_1). The difference of value-judgements is statistically highly significant at a level of $p \leq .01$. This difference is established both by an increased valuation of ideas that address user-needs (here the mean valuation moves up from 0.34 in the pre-test to 0.47 in the post-test) and by a decreased valuation of ideas that would not help the pre-defined target user (average ratings drop from 0.22 in the pre-test to −0.03 in the post-test).

Regarding inclinations to intervene, a similar picture obtains (see Table 5.2). At first, the existing or lacking focus on user needs does not inform the coaches' inclination to intervene in their team's creative process. After the training, the coaches feel more inclined to intervene when their team lacks a focus on user needs

Table 5.1 Perceived value of ideas with vs. without focus on user needs

	Value judgements when ideas	Mean	Std. dev.	t	Sig. (2-tailed)
t_0	Focus on user needs	0.34	0.60	0.88	0.39
	Don't focus on user needs	0.22	0.49		
t_1	Focus on user needs	0.47	0.49	4.29	0.00**
	Don't focus on user needs	−0.03	0.52		

t_0 pre-training assessment, *t_1* post-training assessment, *std. dev.* standard deviation, *sig.* level of significance, t-values of a paired-samples t-test, ** result statistically significant at a level of p≤.01

Table 5.2 Wish to intervene when ideas focus vs. do not focus on user needs

	Wish to intervene when ideas	Mean	Std. dev.	t	Sig. (2-tailed)
t_0	Focus on user needs	−0.01	0.87	0.04	0.97
	Don't focus on user needs	−0.01	0.79		
t_1	Focus on user needs	−0.06	0.62	−1.58	0.13
	Don't focus on user needs	0.13	0.68		

(the intervention mean is positive with a value of 0.13, indicating that on average the coaches want to intervene). Conversely, ideas that attend to user needs incline the coaches to let their teams proceed incessantly, without interventions (here the intervention mean is negative with a mean value of −0.06, indicating that on average the coaches feel rather inclined to not-intervene). However, this difference does not quite reach statistical significance.

Generally, a pattern crystallises that can be observed on all four idea dimensions. To avoid repetitions we will only discuss it once, here pertaining to the focus on user needs. The participants show considerable consensus when providing value estimates for ideas. Standard deviations on value items are comparatively small, even prior to the training. In this case, they amount to .60 and .49 (see t_0 in Table 5.1). By contrast, regarding the question whether or not to intervene in the team's creative process the coaches initially articulate strongly diverging intuitions. In the pre-training assessment, standard deviations on intervention items are rather large (.87 and .79). However, from pre- to post-training the coaches develop more homogeneous intuitions as to when they should intervene. At t_1, standard deviations on intervention items have dropped to .62 and .68.

5.4.2 Balancing Team Interests

Prior to their training, the coaches strongly attend to the factor of team member interests. They significantly prefer ideas, which accommodate the interests of all team members. This preference is statistically highly significant at a level of p≤.01 (see Table 5.3, results at t_0). After their training, the coaches still prefer ideas that balance team interests. However, the impact of this factor has declined; the

Table 5.3 Perceived value of ideas that balance vs. do not balance team interests

	Value judgements when ideas	Mean	Std. dev.	t	Sig. (2-tailed)
t_0	Balance team interests	0.50	0.48	3.46	0.00**
	Don't balance team interests	0.06	0.61		
t_1	Balance team interests	0.30	0.51	1.26	0.22
	Don't balance team interests	0.15	0.52		

Table 5.4 Wish to intervene when ideas balance vs. do not balance team interests

	Wish to intervene when ideas	Mean	Std. dev.	t	Sig. (2-tailed)
t_0	Balance team interests	−0.16	0.91	−2.33	0.03*
	Don't balance team interests	0.18	0.81		
t_1	Balance team interests	−0.06	0.68	−1.26	0.22
	Don't balance team interests	0.10	0.71		

*Result statistically significant at a level of $p \leq .05$.

difference of value-judgements is not statistically significant any more (see Table 5.3, results at t_1).

In the pre-test the intuitions of the coaches as to whether or not they should intervene are also clearly attuned to interests of design team members (see Table 5.4, results at t_0). When solution ideas fail to pick up on some team members' interests the coaches feel inclined to intervene (with a mean of 0.18 tending positively towards interventions). Conversely, the coaches do not intervene but let teams proceed when the pursued solution idea accommodates differing team member interests (in that case, the inclination-mean figures in the negative realm at −0.16).

After the training, coaching strategies are still somewhat attuned to the balancing of team interests, but the factor is less important than it was in the beginning (see Table 5.4, results at t_1). It does not make a statistically significant difference any more whether ideas balance or do not balance team interests.

5.4.3 Originality

Before the training, the coaches do not attend to the factor "originality". The average perceived idea value is the same (.28 and .28, see t_0 in Table 5.5), irrespective of whether ideas are original or unoriginal. After the training, the coaches have become highly sensitive to the originality-dimension of ideas. However, contrary to our hypotheses, the coaches strongly prefer unoriginal ideas (see t_1 in Table 5.5). The average value-estimate of original ideas now figures in the negative realm at −.02, while the average value-rating for unoriginal ideas is even increased compared to the pre-test and now amounts to.46.

The repudiation of originality after the training is in fact a consistent pattern that obtains even on a more fine-grained level of analysis. About the same value-difference in favour of non-originality is found when groups of immediately

Table 5.5 Perceived value of original vs. unoriginal ideas

	Value judgements when ideas	Mean	Std. dev.	t	Sig. (2-tailed)
t_0	Are original	0.28	0.56	0.03	0.98
	Are unoriginal	0.28	0.45		
t_1	Are original	−0.02	0.55	−4.92	0.00**
	Are unoriginal	0.46	0.41		

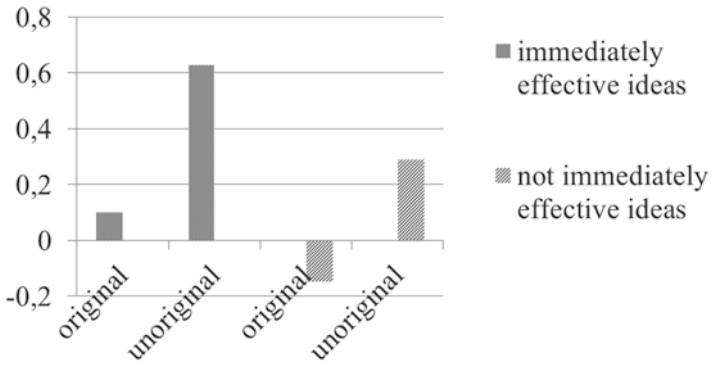

Fig. 5.8 In the post-training assessment, coaches consistently prefer unoriginal ideas. This pattern holds both when they evaluate immediately effective ideas (which on average are attributed greater values, cf. means of .10 and .63) and when they assess not immediately effective ideas (which on average are attributed lesser value, cf. means of −.15 and .29)

Table 5.6 Wish to intervene when ideas are original vs. unoriginal

	Wish to intervene when ideas	Mean	Std. dev.	t	Sig. (2-tailed)
t_0	Are original	−0.01	0.90	−0.04	0.97
	Are unoriginal	−0.01	0.75		
t_1	Are original	0.20	0.66	2.73	0.01**
	Are unoriginal	−0.14	0.66		

effective vs. not immediately effective ideas are analysed separately (see Fig. 5.8). In the discussion we will return to this issue.

As the coaches demonstrate no "antennas" for originality in the pre-test they also disregard this aspect in their coaching at first. The average inclination to intervene is the same (−.01 and − .01, see t_0 in Table 5.6), irrespective of whether teams pursue original or unoriginal ideas. After the training, the coaches have developed highly sensitive antennas for originality and in fact appear to censor ideas with this characteristic. As the positive value of .20 indicates, the coaches want to intervene when their teams pursue original ideas (see t_1 in Table 5.6). By contrast, the negative value of −.14 indicates that teams are left to proceed incessantly when unoriginal ideas are pursued.

Table 5.7 Perceived value of immediately effective vs. not immediately effective ideas

	Value judgements when ideas	Mean	Std. dev.	t	Sig. (2-tailed)
t_0	Are immediately effective	0.71	0.61	6.65	0.00**
	Are not immediately effective	−0.15	0.49		
t_1	Are immediately effective	0.37	0.47	2.90	0.01**
	Are not immediately effective	0.07	0.51		

Table 5.8 Wish to intervene when ideas are immediately effective vs. not immediately effective

	Wish to intervene when ideas	Mean	Std. dev.	t	Sig. (2-tailed)
t_0	Are immediately effective	−0.14	1.02	−1.20	0.24
	Are not immediately effective	0.12	0.84		
t_1	Are immediately effective	−0.05	0.69	−1.17	0.26
	Are not immediately effective	0.11	0.66		

5.4.4 Immediate Effectiveness

Prior to the training, the coaches base their value-judgements clearly on the dimension of immediate effectiveness (see t_0 in Table 5.7). Ideas that lack immediate effectiveness receive a negative average rating of −.15. By contrast, the mean rating of immediately effective ideas is strikingly positive at a value of .71. After the training, immediately effective ideas are still preferred, but the difference is not as large any more (.37 vs. .07 at t_1).

The inclinations to intervene tend to mirror the coaches' preference for immediately effective ideas. The coaches rather want to intervene when teams pursue not-immediately effective solutions (with a mean of .12 at t_0 and .11 at t_1, see Table 5.8). Conversely, the coaches rather refrain from interventions on their team's solution trajectory when the pursued idea seems immediately effective (−.14 at t_0 and − .05 at t_1). However, coaching inclinations do not differ to a statistically significant degree, neither in the pre- nor in the post-test. Thus, a noteworthy discrepancy obtains between the clarity of personal preferences for immediately effective ideas on the one hand and rather indifferent coaching approaches on the other. We will return to this issue in the discussion.

5.4.5 Predicting Value-Attribution Based On Idea Characteristics

Linear regression models are computed to estimate the impact of each factor (focus on user needs, balancing team interests, originality, immediate effectiveness) on the coaches' attribution of value to ideas. We permitted no computation of a regression constant to render the beta weights of the four idea dimensions more easily comparable.

Table 5.9 A linear regression model predicting perceived idea value, before the training

	Unstand. Beta	Standard. Beta	t	sig.
Focus on user needs	−.07	−.04	−.60	.55
Balancing team interests	.27	.15	2.45	.02
Originality	−.17	−.09	−1.53	.13
Immediate effectiveness	.70	.38	6.34	.00

unstand. Beta unstandardized beta coefficient, *standard. Beta* standardized beta coefficient

The two regression models for the pre- and post-test data both explain only a limited amount of variance (R = .41 in the pre-test model and .35 in the post-test model; both models are statistically highly significant at a level of $p \leq .001$). This is in line with prior expectations, since differing personal perspectives on ideas are considered a resource in creative communities and trainings do not serve the purpose of creating uniform responses.

The pattern of significant vs. insignificant beta weights in the regression models mostly accords with the study hypotheses.

When the pre-training data is analysed, the standardized beta coefficient of the factor "focus on user needs" is close to zero and not statistically significant (see Table 5.9). This, again, suggests that the coaches do not attend to this dimension prior to their training.

"Balancing team interests" is the second best predictor for value-ratings of the coaches. The standardized beta weight of this factor amounts to .15, which is statistically significant at a level of $p \leq .05$. This idea dimension clearly informs value-ratings of the coaches.

The beta weight of "originality" is not statistically significant. Prior to their training, the coaches obviously do not screen this idea characteristic systematically in the process of estimating idea value.

"Immediate effectiveness" is the best predictor for value-ratings. The standardized beta weight of .38 is the largest in the whole model, reflecting a strong increase in perceived idea value when ideas are immediately effective. This beta weight is statistically significant at a level of $p \leq .01$.

After the training, the factor "focus on user need" has become the most important predictor for value-judgments (see Table 5.10). Its standardized beta weight of .28 is statistically significant at a level of $p \leq .01$.

The impact of the factor "balancing team interests" has considerably declined. Its beta weight of .08 is not statistically significant any more.

The factor "originality" has become the second best predictor for value estimates. However, the standardized beta weight is negative at a value of −.27, which is statistically highly significant at a level of $p \leq .01$. Thus, increases in idea originality predict reduced value-attributions by the coaches.

The impact of the factor "immediate effectiveness" has dropped considerably. The beta weight now amounts to .16 (compared to .38 in the pre-test), which is, however, still statistically significant on a level of $p \leq .01$. Figure 5.9 summarises the outcomes on behalf of all study hypotheses.

Table 5.10 A linear regression model predicting perceived idea value, after the training

	Unstand. Beta	Standard. Beta	t	sig.
Focus on user needs	.50	.28	4.51	.00
Balancing team interests	.15	.08	1.32	.19
Originality	−.49	−.27	−4.42	.00
Immediate effectivity	.29	.16	2.64	.01

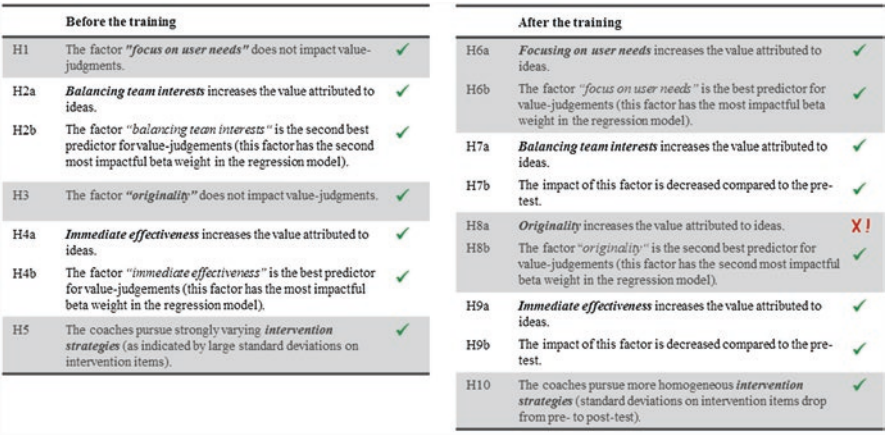

Fig. 5.9 Overview of study hypotheses and respective findings

5.5 Discussion

The evaluation style of coaches who attend the *Design Thinking Certification Program* clearly changes over time.

Prior to their training, the coaches display a static assessment style, evaluating ideas in the middle of the process as though they were facing final products. All ideas should be immediately effective, else wise they are considered poor. Furthermore, design thinking values such as the focus on user needs do not inform idea perceptions of the coaches. Ways of incorporating team-dynamics in the process are also not aligned to design thinking practices.

After the training, the coaches have adopted a dynamic assessment style, de-emphasizing the importance of "immediate effectiveness" in the ideation phase. Their value-set has changed mostly in accordance with design thinking teachings. In particular, coaches now favour ideas that focus on user needs. This factor has become the most important predictor for value judgements. In addition, team-dynamics are still attended, but do not lead to an overly rigorous constriction of the solution space; teams are not expected any more to gratify multiple user needs with their solution simply to accommodate differing team member interests. Finally, the coaches have become highly sensitive to the dimension of idea originality. However,

this idea characteristic factors in negatively in the coaches' estimation of idea value, which is contrary to design thinking teachings and a highly surprising result.

As an overall finding, this study shows that idea assessment styles are not fixed. They are no unalterable personality characteristics. People can learn to perceive ideas in novel ways due to creativity education. Moreover, methodologically it is possible to measure the idea assessment styles of single estimators and groups, and to trace changes in their evaluation styles in the course of creativity trainings.

While the overall results are certainly multi-facetted, we shall confine further discussions to two issues only, which are likely to stimulate further research. One issue concerns a deeper understanding of valuation processes, the other bears on the handling of idea originality.

1. How do emotional and cognitive aspects figure in value estimations, and how do they influence coaching/teaching behaviours?

To assess the attribution of value to ideas, in our study the participants were asked a general question: "How do you find this idea?" Notably, answers to such a general question can reflect both emotional and cognitive aspects of valuation. When a coach states "I find this idea terrible", she may sense a strong emotional aversion while cognitively believing that the idea would work for other people. Or she might consider the idea futile on a cognitive level, while not feeling much about the subject at all. Of course, mixed cases can occur just as well.

One might assume that coaches translate their perceptions of idea value more readily into coaching behaviour when the valuation accords with cognitive reasoning. In our pre-training assessment two strikingly different patterns crystallise on behalf of the factors "balancing team interests" and "immediate effectiveness". In the pre-training assessment both dimensions strongly inform the coaches' attribution of value to ideas. However, only the factor "balancing team interests" has a significant impact on coaching interventions. This pattern makes sense if "balancing team interests" is cognitively construed as advantageous. After all, literature on social competence is often taken to suggest that the balancing of team interests be important and favourable. By contrast, "immediate effectiveness" could be a characteristic the coaches prefer emotionally, but not necessarily on a cognitive level. Probably it feels good when team ideas are expected to work out quickly and smoothly. At the same time, cognitively the coaches might still believe that innovation projects should be open to ideas that lack immediate effectiveness. After all, satellites, GPS and mobile phones – to name just a few examples – certainly did not achieve immediate effectiveness right after the ideas were first conceived. Such a discrepancy between emotional and cognitive appraisals could explain why the coaches value immediately effective ideas more highly, but do not base their coaching strategies consistently on this preference.

If this were true, the *Certification Program* was likely beneficial in helping the coaches emotionally handle ideation outcomes that lack immediate effectiveness. After the training, the coaches' preference for immediately effective ideas is greatly reduced, while the respective coaching strategy remains almost unchanged. This would be an excellent emotional learning outcome for coaches in innovation proj-

ects, where having to handle highly original, but not immediately effective ideas may be everyday business.

In any case, emotional and cognitive aspects of idea valuation can be a fruitful subject for creativity research, well-beyond the training of design thinking coaches. In creativity education, or even education quite generally, teachers have to react – often spontaneously – to student ideas. To support teachers (or design thinking coaches) in this difficult task, a better understanding of the underlying estimation processes would seem very helpful. Research could, for instance, focus on the following set of questions:

- What immediate feelings do specific types of student ideas (e.g., original vs. unoriginal) elicit in teachers/coaches?
- What cognitive rationales do teachers/coaches follow when they decide about behavioural reactions to student ideas?
- What heuristics should teachers/coaches follow when they react to student ideas (e.g., in order to help students develop creative mindsets)?
- How can we train teachers/coaches, so that their immediate emotional reactions to student ideas facilitate favourable behavioural reactions?

Studies bearing on these or related issues have already been undertaken in different contexts and much progress can be expected from a knowledge synthesis across different domains (e.g., Zajonc 2001; Oreg 2006; Beghetto 2016; Corazza 2017). We are curious what the community of creativity researchers will jointly bring to light over time.

The second study finding to be discussed in more detail concerns specifically the handling of originality.

2. *Why do the study participants develop disdain for original ideas?*

In our longitudinal study, coaches strongly repudiate original ideas in the post-training assessment. This is an unexpected, yet statistically highly significant outcome. We shall consider a number of potential explanations in turn.

As has been noted above (in Sect. 5.2.3), original ideas often necessitate prolonged periods of refinement to render them effective. Might the coaches have sensed a conflict between the higher resource demand in the case of original ideas and limited available project resources, such as remaining time? That could explain why they would prefer unoriginal ideas. However, for several reasons this explanation does not seem to work in the case of the present study. First, the coaches answered questions on behalf of a purely hypothetical scenario where no information was even mentioned about available project resources. Second, the ideas had been artificially generated so as to not entail different resource demands. On average the original ideas were just as immediately effective (easy to implement and likely to produce the intended effects, judged by two expert raters) as were the unoriginal ideas. Thirdly, if the study participants had sensed a conflict between idea originality and available project resources, this conflict should have existed in the pre- and the post-test alike. Yet, only in the post-test did the coaches display a significant preference for unoriginal ideas. Finally, even if we assume that the

coaches saw conflicts between idea originality and available project resources in the post-test only, they could still have personally liked the original ideas. In that case, they might have stated in the questionnaire that they personally found the original ideas excellent, while at the same time they would have launched coaching interventions against them in light of limited project resources. However, this pattern was not observed, quite to the contrary. The coaches were much clearer in devaluing original ideas ($t = -4{,}92$) than they were in launching coaching interventions when faced with original ideas ($t = 2.73$). Here t-values that diverge more greatly from zero indicate stronger differences between original and unoriginal ideas.

Considering further explanations proposed in the literature for why people might prefer unoriginal ideas, we can also return to Blair and Mumford (2007). They emphasize how it is more difficult for audiences to forecast the effects of original ideas, since these are unprecedented. Hurdles might emerge in the implementation phase and social benefits might be less foreseeable. Again, however, in our study the ideas were artificially designed to avoid spurious correlations between idea originality and (unclear) implementation difficulties or (unclear) social benefits. On average, the original and unoriginal ideas did not differ in their immediate effectiveness or their social favourability/serviceability to user needs. Notably, it is also quite easy to come up with unoriginal ideas that entail great forecasting difficulties. For instance, in the case of the unoriginal idea mentioned in Sect. 5.3.1, the plan is to make a farmer happy by finding a spouse for him via a partnership ad. It is hard to foretell what hurdles will arise given this solution strategy and how the social benefits or misfortunes will balance out in the end. Different women answering the ad could affect the farmer's happiness in various ways. It is unclear whether a person who publishes a partnership ad will ever find a spouse, let alone one who makes him happy and helps him at work. This is not to deny that particularly great forecasting difficulties can explain the disdain for original ideas in some contexts, but for the present study results a different explanation seems needed.

Barron's (1955) research suggests that people display dispositions towards originality: Some people favour the original, other people favour the conventional consistently across different situations. For Barron, these dispositions develop throughout childhood and they become so stable that they can be addressed as personality traits in adult populations. We only reported group results above. They did, however, also indicate something in the direction of cross-situational dispositions. The coaches consistently favour unoriginal ideas in the post-training assessment. More specifically, this preference becomes evident in two different situations, namely when the assessed ideas are immediately effective and when the ideas lack immediate effectiveness (cf. Fig. 5.8). Notably, though, these cross-situational dispositions to prefer the unoriginal are not carved in stone. Preference patterns change from the pre- to the post-test. Thus, our study results are compatible with the belief that people develop particular dispositions to either prefer the original or the conventional across different situations. Importantly, though, these dispositions do not seem to be unalterable personality traits. They can change in the course of creativity trainings. While in the case of our study the coaches, unexpectedly and against the

trainers' intentions, developed a disposition to favour the conventional, an important point is that the dispositions did change.

The question remains why the participants in our study developed disdain for original ideas. Some qualitative feedback provided by the study participants after their training gave hints to make sense of the findings. As a number of coaches indicated in personal conversations about the study outcomes, they had picked up the importance design thinking experts attributed to originality in the course of their training, as for instance the motto "encourage wild ideas" was placed centre-stage in ideation sessions. Still, the coaches were not convinced that wild ideas would eventually entail better project results. Some coaches also indicated that they found "crazy-sounding ideas unsuitable and not feasible in the more serious work contexts" where they lead innovation teams. Follow-up qualitative research needs to clarify these issues more systematically. Nonetheless, an important objective for creativity educators becomes apparent, which is likely relevant beyond design thinking trainings: *Creativity education needs to show how originality leads to something better, not just something different*.

Indeed, even in our test scenario originality was not associated with *better* ideation outcomes. As our questionnaire had been designed like this, original ideas were not associated with greater immediate effectiveness or clearer social benefits/more gratification of user needs. Had the coaches preferred original ideas, they would have preferred originality for its own sake, as an idle idea characteristic, unrelated to idea effectiveness. The coaches did not appreciate originality for its own sake, and there may even be good reasons for such a stance as long as "radical innovation" or "creativity" are no self-standing goals. After all, in real-world challenges original ideas often do necessitate a greater investment of resources later on, e.g., more time is required to refine the idea, eliminate bugs and make the approach work (cf. Sect. 5.2.3). Moreover, as Blair and Mumford (2007) point out, potential positive or negative effects of unprecedented, original ideas can be difficult to anticipate.

All in all, the handling of originality appears to be a particularly delicate learning and teaching objective. In our study everything else seemed easier. The coaches readily de-creased their preference for immediately effective solutions. They readily attended to a novel characteristic, the focus on user needs, making it their primary criterion for idea value. They readily changed the handling of team dynamics. However, the coaches did not readily accept standard teachings on how to handle originality.

Yet, originality is a defining characteristic of creative achievements. Therefore, it seems an important objective for creativity education to help learners develop proficiency in the handling of original ideas. A consistent experience of disdain for original ideas is likely to be a major creativity block that research and education need to tackle. Against this background, we will close this chapter with considerations how the handling of originality might be better facilitated in the future.

5.6 Outlook: How to Facilitate the Acceptance of Originality?

A creative mindset must include some openness to original ideas. While it may be unnecessary (or even unfavourable) that people prefer all novel ideas simply for the sake of originality, people who aspire to be creative, or who want to lead innovation teams, must be ready to embrace at least some (promising) novel ideas. Otherwise their projects remain tied to the realm of the conventional and they cannot possibly succeed.

Especially in educational contexts it has already been noted how unconstrained originality – permitting all kinds of novelty – is perceived as daunting and something to be carefully avoided (Beghetto 2016). A likely follow-up question is how originality can be directed along task-appropriate lines. Creativity researchers could advance theories and methods to help learners explain, predict and control the development of creative potential in their projects, based on how they make use of originality in the process. In this context, it seems promising to elucidate the varying roles of originality in different project stages. Moreover, explanations can be sought as to why some original contributions seem much more promising than others in each project stage.

Clearly there are many ways to incorporate originality in a creative process, and not all of them seem equally helpful to build up creative potential. In design thinking projects, as probably in many other creative endeavours, people start by understanding and observing a subject domain. This is already the first phase in which a creative project can be imbued with more or less originality. In this phase, originality could be intentionally induced (i) by means of random imaginings about the field, (ii) by building on any arbitrary observation or (iii) by attending to a surprising, unexpected observation in the field. With these three approaches, the same level of originality might be achieved, but the creative potential that is added to the project does not seem the same. Similar things can be said basically about all steps and means of creative work.

To consider one more sample project phase, we can turn to ideation, i.e. the stage of thinking up ideas and selecting one or more of them to be further pursued. In this phase, not all original ("wild") ideas seem similarly promising. A well-known positive example of "wild ideas" benefitted Polaroid developments. When Edwin Land of Polaroid Corporation first imagined a printed colour picture that would be available within a few seconds after a camera had captured the image, his idea seemed "wild" to the contemporaries: original, bold, almost unrealistic. Land had to work for many years on his project to render the idea effective. Yet, eventually this idea helped to expand the realm of what humanity could do (Arnold 1959/2016). His wild idea had endowed Land's project with a great creative potential. In other cases, wild ideas may be truly impossible to realise. They could still be helpful, e.g., as metaphors that guide the search for feasible solutions. By contrast, other original/wild ideas may do very little in terms of adding creative potential. It is not the wildness of ideas per se that gets a project far. In design thinking, examples of original/

wild ideas that do not seem particularly promising can be easily generated by imagining a new, unrealizable solution that would not even make the user happy if it could be implemented.

When estimators dislike originality because they do not see how it improves overall project outcomes, they have a point to make. Not all originality is necessarily productive. Moreover, we know that originality often comes at a cost. However, when creativity educators can explain and demonstrate how to use originality in beneficial ways, the estimators we train have no reason to feel badly about it in lump-sum ways.

We submit the following questions for further discussions:

- How do learners experience the role of originality in their processes? Do they experience originality as a route to something better, or merely as a route to something different? How are potential negative effects of originality construed (such as increased resource demands)?
- What techniques of producing and handling originality help to amplify a project's creative potential? Are different techniques required in different work stages?
- What distinguishes helpful original contributions in a creative process from less helpful ones? Or can all original contributions endow projects with a high creative potential if only the original elements are processed in a particular way?
- How does original information build up in the course of creative work? (E.g., what is the relationship between original notes about a research domain, developments of original problem views, novel solution ideas etc.?)
- How do imagination abilities impact the acceptance of original ideas? Can people picture conventional solutions better than original ones? Might people simply prefer solutions they can easily imagine?[1]
- Extensive literature treats the subject of resistance to innovation and change (e.g., Oreg 2006; Talke and Heidenreich 2013). To what extent do the phenomena of people repudiating change/innovation and people repudiating originality overlap or differ from one another?
- How does the acceptance of original ideas differ in cases when they have been (a) thought up by oneself, (b) developed in one's own creative team, (c) proposed by a friend or (d) by someone else?

All in all, it seems the process of learning how to judge creative work is dynamic in a double sense. Single estimators learn novel assessment styles in the course of creativity trainings. At the same time, creativity experts are still in the process of finding out which assessment styles to recommend and how to convey them. In any case, people's intuitions concerning idea value do not seem to be carved in stone. They rather appear to be readily changeable by means of trainings. This, of course, entails great responsibilities on the part of creativity researchers, whose views of the creative process impact the way in which teachers and coaches assess student ideas.

[1] These research questions were suggested to us by Axel Menning.

Acknowledgements We thank Ronald Beghetto and Giovanni Emanuele Corazza for inviting this research to be part of their comprehensive project on dynamic perspectives in creativity education and helpful feedback on our first draft. We thank the participants of the *Certification Program for Design Thinking Coaches* in the summer semester of 2017 at the *HPI Potsdam* for taking part in our survey and for their openness to subsequent individual discussions of study results. We thank Toni Mattis for the permission to print his picture.

References

Agnoli, S., Mastria, S., Kirsch, C., & Corazza G. E. (under review). *Creativity in the advertisement domain: The mediating role of experience on creative achievement*.

Arnold, J. E. (2016). Creative engineering. In W. J. Clancey (Ed.) *Creative engineering: Promoting innovation by thinking differently* (pp. 59–150). Stanford : Stanford Digital Repository. Original manuscript 1959. http://purl.stanford.edu/jb100vs5745. Accessed 13 June 2017.

Barron, F. (1955). The disposition toward originality. *Journal of Abnormal and Social Psychology, 51*(3), 478–485.

Beaney, M. (2005). *Imagination and creativity*. Milton Keynes: Open University.

Beghetto, R. A. (2007). Does creativity have a place in classroom discussions? Prospective teachers' response preferences. *Thinking Skills and Creativity, 2*(1), 1–9.

Beghetto, R. A. (2010). Creativity in the classroom. In J. C. Kaufman & R. J. Sternberg (Eds.), *Cambridge handbook of creativity*. New York: Cambridge University Press.

Beghetto, R. A. (2013). Nurturing creativity in the micro-moments of the classroom. In K. H. Kim, J. C. Kaufman, J. Baer, & B. Sriraman (Eds.), *Creatively gifted students are not like other gifted students. Advances in creativity and giftedness* (Vol. 5). Rotterdam: Sense Publishers.

Beghetto, R. A. (2016). Learning as a creative act. In T. Kettler (Ed.), *Modern curriculum for gifted and advanced learners* (pp. 111–127). New York: Routledge.

Blair, C. S., & Mumford, M. D. (2007). Errors in idea evaluation: Preference for the unoriginal? *Journal of Creative Behavior, 41*(3), 197–222.

Brown, T., & Katz, B. (2009). *Change by design: How design thinking transforms organizations and inspires innovation*. New York: Harper Collins.

Carleton, T., & Leifer, L. (2009). Stanford's ME310 course as an evolution of engineering design. In R. Roy & E. Shehab (Eds.), *Proceedings of the 19th CIRP design conference – Competitive design* (pp. 547–554). Cranfield: Cranfield University.

Cojuharenco, I., Cornelissen, G., & Karelaia, N. (2016). Yes, I can: Feeling connected to others increases perceived effectiveness and socially responsible behavior. *Journal of Environmental Psychology, 48*, 75–86.

Corazza, G. E. (2015). *Impresa e creatività*. Online lecture. https://www.youtube.com/watch?v=mfR8qJWtpfs. Accessed 14 Sept 2017.

Corazza, G. E. (2016a). Potential originality and effectiveness: The dynamic definition of creativity. *Creativity Research Journal, 28*(3), 258–267.

Corazza, G. E. (2016b). *Creativity: A dynamic definition*. Keynote speech at the MIC Conference 2016: From creative brains to creative societies, Università di Bologna, Bologna, 14–16 September 2016**.**

Corazza, G. E. (2017). *Creativity in education: a recursive exercise in estimation ability*. Keynote speech at the Utrecht platform for creativity in education, Ut-recht University, Utrecht, 30–31 March 2017. https://platformcreativity.sites.uu.nl/wp-content/uploads/sites/145/2017/04/2017-UPCE-Giovanni-E.-Corazza.pdf. Accessed 14 Oct 2017.

d.school. (2010). *Bootcamp bootleg*. http://dschool.stanford.edu/wp-content/uploads/2011/03/BootcampBootleg2010v2SLIM.pdf. Accessed 6 Mar 2017.

Dow, S. P., & Klemmer, S. R. (2011). The efficacy of prototyping under time constraints. In H. Plattner, C. Meinel, & L. Leifer (Eds.), *Design thinking research. Understand – improve – apply* (pp. 111–128). Heidelberg: Springer.

Gaut, B. (2010). The philosophy of creativity. *Philosophy Compass, 5*(12), 1034–1046.
Greenberg, R. (2003). The role of neophobia and neophilia in the development of innovative behaviour of birds. In S. M. Reader & K. N. Laland (Eds.), *Animal innovation* (pp. 175–196). Oxford: Oxford University Press.
Hasso Plattner Institute for Digital Engineering. (2017). *This is design thinking: A collection of case stories*. http://thisisdesignthinking.net. Accessed 29 Oct 2017.
Kaufman, A. B., & Kaufman, J. C. (2014). Applying theoretical models on human creativity to animal studies. *Animal Behavior and Cognition, 1*(1), 78–90.
Kelley, T., & Kelley, D. (2013). *Creative confidence*. London: Harper Collins.
McKim, R. H. (1972). *Experiences in visual thinking*. Belmont: Wadsworth Publishing.
McKim, R. H. (2016). Designing for the whole man. In W. J. Clancey (Ed.) *Creative engineering: Promoting innovation by thinking differently* (pp. 198–217). Stanford: Stanford Digital Repository. Original manuscript 1959. http://purl.stanford.edu/jb100vs5745. Accessed 13 June 2017.
Ney, S. (2016). *Manual for design thinking coaches*. Potsdam: HPI School of Design Thinking.
Oreg, S. (2006). Personality, context and resistance to organizational change. *European Journal of Work and Organizational Psychology, 15*(1), 73–101.
Roth, B. (2015). *The achievement habit*. New York: Harper Collins.
Runco, M. A., & Jaeger, G. J. (2012). The standard definition of creativity. *Creativity Research Journal, 24*(1), 92–96.
Talke, K., & Heidenreich, S. (2013). How to overcome pro-change bias: Incorporating passive and active innovation resistance in innovation decision models. *Journal of Product Innovation Management, 31*(5), 894–907.
von Thienen, J. P. A., Noweski, C., Meinel, C., & Rauth, I. (2011). The co-evolution of theory and practice in design thinking – or – "mind the oddness trap!". In H. Plattner, C. Meinel, & L. Leifer (Eds.), *Design thinking research. Understand – improve – apply* (pp. 81–99). Heidelberg: Springer.
von Thienen, J. P. A., Royalty, A., & Meinel, C. (2016). Design thinking in higher education: How students become dedicated creative problem solvers. In C. Zhou (Ed.), *Handbook of research on creative problem-solving skill development in higher education* (pp. 306–328). Hershey: IGI Global.
von Thienen, J. P. A., Clancey, W. J., Corazza, G. E., & Meinel, C. (2017a). Theoretical foundations of design thinking. Part I: John E. Arnold's creative thinking theories. In H. Plattner, C. Meinel, & L. Leifer (Eds.), *Design thinking research. Making distinctions: Collaboration versus cooperation* (pp. 13–40). Cham: Springer.
von Thienen, J. P. A., Paladini, C., & Meinel, C. (2017b). *Do creative ideas give rise to unique private-mind or private-language problems?* Paper presented at the 25th annual conference of the European Society for Philosophy and Psychology, University of Hertfordshire, Hertfordshire, 14–17August 2017.
von Thienen, J. P. A., Clancey, W. J., & Meinel, C. (forthcoming-a). Theoretical foundations of design thinking. Part II: Robert McKim's need based design theory. In H. Plattner, C. Meinel, & L. Leifer (Eds.), *Design thinking research*. Cham: Springer.
von Thienen, J. P. A, Ney S., Meinel C. (forthcoming-b). *Idea Assessment Probes. Questionnaires to measure the idea assessment style of estimators in creative projects*. Electronic Colloquium on Design Thinking Research.
Zajonc, R. B. (2001). Mere exposure: A gateway to the subliminal. *Current Directions in Psychological Science, 10*(6), 224–228.