



Universiteit
Leiden
The Netherlands

creatIef, Creatiever, creaTiefst? Onderzoek naar het belang en de werking van creativiteitstraining bij Bachelor of ICT-studenten
Blok, B.Z.

Citation

Blok, B. Z. (2020, December 1). *creatIef, Creatiever, creaTiefst? Onderzoek naar het belang en de werking van creativiteitstraining bij Bachelor of ICT-studenten*. SIKS Dissertation Series. Creativity & Innovation Foundation. Retrieved from <https://hdl.handle.net/1887/138481>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/138481>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/138481> holds various files of this Leiden University dissertation.

Author: Blok, B.Z.

Title: creatIef, Creatiever, creaTiefst? Onderzoek naar het belang en de werking van creativiteitstraining bij Bachelor of ICT-studenten

Issue date: 2020-12-01

13 Praktische-relevantiestudie alsmede het gezamenlijk effect van creativiteitstraining

In dit hoofdstuk beschrijf ik de samenvoeging (onderzoekssynthese) van de resultaten – met de nadruk op de effectsterkten – uit de drie studies van dit meervoudige empirisch onderzoek (zie § 8.4, voor de introductie daarvan). De *Praktische-relevantiestudie* dient om antwoorden te vinden op de tweede onderzoeksvraag: “Werkt creativiteitstraining die studenten *Informatie- en Communicatietechnologie* van Nederlandse hogescholen volgen vanuit het curriculum?”

De hiervoor omschreven synthese betreft de bevindingen van de Interventiestudie (*Hoofdstuk 10*), de Enquêtestudie (*Hoofdstuk 11*), en de Experts-studie (*Hoofdstuk 12*). Middels die samenvoeging van resultaten ontstaat er een totaalbeeld vanuit een kwalitatief perspectief. De specifieke vraag bij de Praktische-relevantiestudie luidt: “Wat is de gezamenlijk effectomvang en de zogeheten praktisch relevantie van de werking van creativiteitstraining bij Bachelor of ICT-studenten?”

Dit hoofdstuk is verder opgebouwd uit acht secties: *Praktische significantie of praktische relevantie?* (§ 13.1); *Controle van effectgrootten* (§ 13.2); *Gecombineerde effectomvang & praktische relevantie* (§ 13.3); *Bevindingen & discussies* (§ 13.4); *Hoofdstukconclusie* (§ 13.5); *Antwoord op de tweede onderzoeksvraag* (§ 13.6); *Overzicht van alle hoofdstukconclusies van Deel II* (§ 13.8); en *Tot slot van de tweede onderzoeksvraag* (§ 13.8).

13.1 Praktische significantie of praktische relevantie?

Als er onderzoek gedaan wordt naar het effect van iets op iets – in dit geval de werking van een training bij studenten – dan is het van groot belang inzicht te krijgen in de sterkte van dat effect. De traditionele *statistische significantie* lijkt daarvoor niet afdoende.²⁷³ In dat opzicht spreekt het citaat van Morris & Fritz (2013, p. 580) voor zich:

Psychology research is still dominated by null hypothesis significance testing (NHST). However, almost any effect will be significant if a very large sample is tested – something that has become easier to accomplish with the internet and other modern technology. Conversely, it is sometimes only possible to test a small number of participants from, for example, a special population. In that case, even a substantial difference may not be significant. NHST is dependent upon the size of the sample that is tested and the result gives only a binary ‘Yes/No’ decision. Normally, our real interest is in the size of the effect. With a very large sample we might find a statistically significant effect but decide that it was too small for our further attention. Conversely, a large effect from a special, small population, even though not significant, would encourage us to seek ways of testing more participants and combining the data.

273 (O.a.: Cumming, 2012; Hooren van et al., 2015)

274 (O.a.: Cohen J., 1994; Cooper H.M. et al., 1994; Cumming, 2012; Ellis, 2010; Glaser, 1999; Rubin, 2010)

Deze sectie bevat de volgende vier onderwerpen: *Statistische significantie & p-waarde* (§ 13.1.1); *Statistische significantie is géén effectgrootte* (§ 13.1.2); *Effectgrootte & het belang ervan* (§ 13.1.3); en *Sectiesamenvatting & -conclusie* (§ 13.1.4).

13.1.1 Statistische significantie & p-waarde

“Statistical significance does not measure importance” (Kalinowski et al., 2010, p. 51). De aanduiding voor statistische significantie is de zogenoemde *p*-waarde. “The *p*-values characterize only statistical significance, which bears no necessary relationship to practical significance or even to the statistical magnitude of the effect” (Lipsey et al., 2012, p. 3). Statistische significantie wordt dus uitgedrukt in *p*-waarden, waarbij *p* staat voor *probability* (in het Nederlands: *probabiliteit*, *waarschijnlijkheid* of *kans*). De *p*-waarde is een mate van waarschijnlijkheid, “specifically the probability of incorrectly rejecting the null hypothesis when in fact it is true. In other words, it is the probability of a false-positive”, aldus Kalinowski & Fidler (2010, p. 50).

Volgens Field (2009, p. 53), daarentegen, is de consensus dat de experimentele hypothese *waar* is bij een *probability* die kleiner is dan de gestelde drempel *p*-waarde, doorgaans $\leq .05$ of $\leq .01$. Oftewel: er is dan sprake van een effect. Het proces van statische-significantietests is evenwel *dichotoom*, waarbij de uitkomst telkens een keuze is uit twee categorieën, bijvoorbeeld: *man* of *vrouw* (ook *ja-of-nee variabele* genoemd).²⁷⁴

13.1.2 Statistische significantie is géén effectgrootte

De statistische significantie – de *p(robability)*-waarde ofwel *waarschijnlijkheid* - van een meting zegt niets over de sterkte van de uitwerking van een interventie (Kalinowski et al., 2010, p. 50): “A *p*-value is a very limited piece of information, relating to false-positive error rates only. (...) There is much more to know about a set of empirical data!”

Rubin schrijft (2010): “*Statistical significant* does not necessarily mean strong or important, and relationships that are strong or important are not always statistically significant” (p. 150). Ook Glaser (1999) bevestigt dit: “*P*-values say nothing about magnitude or import[ance]” (p. 293); en noemt het *bedrieglijk* en *misleidend* wanneer *p*-waarden in die zin gebruikt worden bij conclusies (p. 294).

Geheel in deze lijn stelde Thompson (2000, p. 1): “*P* values are not useful indices of study effects.” Dit is uiteraard een interessant en zeer relevant gegeven voor de empirische studies in dit proefschriftonderzoek. Boven-

dien merkt Coe (2004, p. 83) nog op: “One of the most telling criticisms of the use of significance tests is that they leave out the most important information: the size of the effect.”

13.1.3 Effectgrootte & het belang daarvan

De grootte van het effect van iets blijkt echter wel degelijk statistisch controleerbaar, zoals al eerder werd geïntroduceerd (§ 8.2.3). Zo is effectgrootte (*effect size: ES*) een objectieve en gestandaardiseerde maateenheid voor de omvang, sterkte of grootte van het geobserveerde effect (Field, 2005); *ES* levert een objectieve meting van de relevantie van een effect van iets. Volgens Ellis (2010, p. 11) is dat een pré: “One advantage of reporting effect sizes in standardized terms is that the results are scale-free, meaning they can be compared across studies. If two studies independently report effects of size $d_{\text{unb}} = .50$, then their effects are identical in size.” Zoals al eerder is genoemd (o.a. in § 8.2), is effectgrootte onafhankelijk van de steekproefomvang, in tegenstelling tot statistische significantie. Daarmee is effectgrootte (*ES*) een betrouwbare maat voor het controleren van het effect van iets bij relatief kleine groepen.

ES is bijvoorbeeld zelfs inzetbaar bij een steekproefomvang van één. Dat maakt het gebruik van effectgrootten (zeer) geschikt voor het onderwijs, omdat verschillende individuele leerprestaties (het effect van lessen en trainingen bij een leerling of student) daarbij een gegeven zijn. Niet voor niets is in het onderwijs de beoordeling van individuele prestaties - anders gezegd: het houden van zeer kleine steekproeven - over het algemeen de norm. Dat is trouwens niet exclusief voor het onderwijs. Zo kunnen bijvoorbeeld in de geneeskunde statistisch relatief zeer kleine, en ook individuele resultaten eveneens van (groot, c.q. zelfs van levens-) belang zijn.²⁷⁵ Effectgrootten worden zelfs gezien als belangrijkste uitkomst van empirisch onderzoek, aldus Lakens (2013, p. 1):

Effect sizes are the most important outcome of empirical studies. Most articles on effect sizes highlight their importance to communicate the practical significance of results. For scientists themselves, effect sizes are most useful because they facilitate cumulative science. Effect sizes can be used to determine the sample size for follow-up studies, or examining effects across studies. (...) Researchers want to know whether an intervention or experimental manipulation has an effect greater than zero, or (when it is obvious an effect exists) how big the effect is.

Daarnaast zegt Cohen (1990, p. 1310) over het belang van effectgrootte (*effect size: ES*): “I have learned and taught that the primary product of a research inquiry is one or more measures of effect size, not *p*-values.” Recentelijker is ook Cumming (2012) duidelijk over de relevantie van effectgrootte: “Estimated *ES*s are usually the main results of research, and should be the main focus of interpretation because they are the best information we have about the population” (p. 40). Tevens bekritiseert Cumming het volgens hem bestaande paradigma dat statistische significantie heilig is.

I included effect sizes and confidence intervals, but the editor tells me there's no room, and I have to take them out and only report *p*-values.

275 (O.a.: Bekkering et al., 2009; Rovers et al., 2012; Rubin, 2010, p. 143)

What should I do? I've been asked that question many times. I hope such questions will quickly fade into history, but we all have to live in the research world as we find it. You may need to write a report or dissertation that pleases the examiners, and we all need to find good journals that will publish our papers. My reply is that the justification for the new statistics is strong, the world should change and is changing, and it's important to keep up our efforts to help it change further. (...) Our most persuasive reason using the new statistics may be that it's simply more informative - it gives a more complete picture of what our data are able to reveal (Cumming, 2012, p. 431).

Effectgrootte & de American Psychological Association

De *American Psychological Association* (APA, 2010) maakt de noodzaak van het rapporteren van effectgrootten als volgt duidelijk: “For the reader to appreciate the magnitude or importance of a study's findings, it is almost always necessary to include some measure of effect size in the Results section” (p. 34). Daarnaast benadrukt APA (p. 33) dat testen op statistische significantie wel een uitgangspunt is, maar dat uitvoerige toelichting en aanvullende statistische elementen, zoals effectgrootte, noodzakelijk zijn bij de verslaglegging om de resultaten zo compleet mogelijk duidelijk te maken.

Om een zo eerlijk en compleet mogelijke voorstelling van de werkelijkheid te geven moet men, zo stelt Cumming (2012), bij studies óók tegenvallende resultaten rapporteren. Dat wordt volgens hem lang niet altijd gedaan. Hij noemt dit het “file drawer effect, (...) the tendency for results that are not statistically significant to remain published. This can seriously bias future meta-analysis.” Cumming (2012, p. 191) haalt aan dat APA (2010) het probleem herkent: “Mention all relevant results ... be sure to include small effect sizes (or statistically nonsignificant findings) ...” (p. 32).”

Doelstelling van effectgrootte

Het gebruik van effectgrootten moet wetenschappers stimuleren, aldus Levine (2010), meer na te denken over de praktische relevantie van hun statistische effecten, in plaats van in hoge mate te vertrouwen op automatismen, zoals het verwerpen of niet van de nulhypothese. Onderzoekers gebruiken effectgrootte dus, naast statistische significantie, voor het inzichtelijk maken van zogeheten *praktische significantie* (o.a.: Sullivan, 2009, pp. 170, 489) of *substantieve significantie* (Ellis, 2010). Sander (2004) verwoordde, bij zijn afscheidscollege aan de *Technische Universiteit Eindhoven*, de toepassing van statistische en praktische significantie als volgt.

Onderzoekers stellen gewoonlijk hypotheses op die met behulp van statistische methoden worden geverifieerd. Daarbij volgt men trouw een reken-schema en concludeert aan het einde dat de uitkomst al dan niet statistisch niet-significant is, en als gevolg hiervan wordt een hypothese wel of niet verworpen. Menig onderzoeker vergeet daarbij dat er twee soorten significantie bestaan: statistische significantie en praktische significantie; en deze twee hebben niets met elkaar te maken. Elk verschil, hoe klein ook, is statistisch significant als er voldoende waarnemingen worden gedaan. In de praktijk is een verschil echter alleen een verschil als het verschil maakt (p. 15).

13.1.4 Sectiesamenvatting & -conclusie

De constatering uit het voorgaande in deze sectie (§ 13.1) is dat er bij (empirisch) onderzoek sprake is van *statistische significantie* en *praktische significantie*. Beide begrippen hebben een andere betekenis, waardoor het woord *significantie* kan leiden tot verwarring, misleiding en het maken van vergissingen.²⁷⁶ Om verwarring tegen te gaan, wordt - conform Cumming (2012) - vanaf dit punt zoveel mogelijk *praktische relevantie* gebruikt in plaats van *praktische significantie*. Cumming (2012, p. 50) adviseert het woord *significant* zeer zorgvuldig te gebruiken, of zelfs te vermijden: “Keep the distinction carefully in mind anyway. Beware the fallacy of the slippery slope of significance.”

Samengevat, staat *statistisch significant* voor (zo goed als zeker) weten dat er effect optreedt; en staat *praktisch relevant* voor een gemeten verschil dat in de praktijk van betekenis is. “It allows us to move beyond the simplistic, ‘does it work or not?’ to the far more sophisticated, ‘how well does it work in a range of contexts?’” (Coe, 2004, p. 80). *Praktische relevantie* en de interpretatie van effectgrootten zijn contextafhankelijk.²⁷⁷ Voorts is effectgrootte een objectieve, gestandaardiseerde maateenheid, die de mogelijkheid biedt om (a) het effect van iets te controleren, en (b) om de resultaten van verschillende studies samen te voegen (te synthetiseren); een en ander onafhankelijk van de steekproefomvang. Oftewel: effectgrootte is ook geschikt voor het controleren van een effect van iets bij relatief (zeer) kleine groepen - waarvan sprake is binnen dit onderzoek. Een dergelijke kennissynthese staat bekend als *meta-analyse*.²⁷⁸ De woorden van Coe (2002, p. 9; 2004, p. 86) bevestigen tot slot het voorgaande: “An important consequence of the capacity of meta-analysis to combine results is that even small studies can make a significant contribution to knowledge.”

13.2 Controle van effectgrootten

Controleren van effectgrootte kan op verschillende manieren. Daarvoor zijn het meest gebruikelijk: Cohen's *d*, *Pearson's correlation coefficient r* en *determinatiecoëfficiënt r²*.²⁷⁹ In onder meer § 8.2 is toegelicht dat bij het huidige onderzoek is uitgegaan van Cohen's *d* - of preciezer: Cumming's d_{unb}^* - voor het bepalen van de sterkte van het effect, om daarmee de resultaten van dit onderzoek te kunnen vergelijken met de effectgrootten van met (meta-analytische) studies van derden, omdat de betreffende auteurs daar gebruik van maakten.

276 (O.a.: Cumming, 2012; Ellis, 2010; Grissom et al., 2005, 2012; Hooren van et al., 2015; Kalinowski et al., 2010; Matheson, 2008; Thompson B., 2002)

277 (O.a.: Coe, 2002; Cohen L. et al., 2007; Cooper H.M. et al., 1994, 2009; Ellis, 2010; Lenth, 2001; Rocconi et al., 2015; Schagen I. et al., 2004)

278 (O.a.: Cooper H.M. et al., 2009; Cumming, 2012; Glass G.V. et al., 1981; Hedges et al., 1985; Rose et al., 1984; Wolf, 1986)

279 (O.a.: Cooper H.M. et al., 1994; Field, 2005; Rubin, 2010)

280 (O.a.: Baguley, 2009; Ellis, 2010; Glass G.V. et al., 1981; Kelley et al., 2012; KNAW, 2009; Lipsey et al., 2012; Magnusson, 2014; Neill, 2006; Schagen I. et al., 2004)

281 (O.a.: Coe, 2004; Driessen, 2007; Driessen et al., 2010; Strand et al., 2006)

Zoals ook bij § 8.2 al is opgemerkt, kan Cohen's *d* een vertekend beeld geven bij (relatief) kleine steekproeven. Hedges (1985) heeft evenwel een manier bedacht om deze bias bij benadering op te heffen. Het is een *kleine* toevoeging op de Cohen's *d* berekening. De uitkomst noemt men Hedges' *g*. *Alhoewel Cumming (2012, p. 294) het d_{unb} (“dee-un-bee”) noemt, van “unbiased”. Hij vindt dat duidelijker, omdat er verwarring kan ontstaan door de vele termen die in omloop zijn en de inconsistentie van het gebruik ervan. Sterker, Cumming schrijft: “The terminology is a mess” (p. 295).

Omdat er bij dit onderzoek sprake is van relatief kleine steekproeven, is Cumming's d_{unb} (oftewel: Hedges' *g*) in de resultaten opgenomen. De biascorrectie is met name sterk bij een steekproefomvang < ongeveer 20. Daarboven neemt het verschil af. Bij een steekproef > pakweg 50 is er nauwelijks meer verschil, althans *niet zorgwekkend* (Cumming, p. 295). Voor de berekening van Cumming's d_{unb} is de *Effect Size Calculator* van Coe (2000a) gebruikt.

Hierna vervolg ik deze sectie met de volgende vier onderwerpen: *Interpretatie van effectgrootte* (§ 13.2.1); *Praktische relevantie bij educatieve interventies* (§ 13.2.2); *Interpretatie van effectgrootte & contextuele aspecten* (§ 13.2.3); en *Sectiesamenvatting* (§ 13.2.4).

13.2.1 Interpretatie van effectgrootte

Een vuistregel is dat bij een effectgrootte (Cohen's *d*) van .20 sprake is van een *klein* effect; vanaf $d = .50$ van een *middelgroot* effect; en bij $d = .80$ en meer van een *groot* effect. Deze richtlijn wordt veel toegepast. Desondanks blijkt deze vuistregel en dus het interpreteren van effectgrootte arbitrair.²⁸⁰ “We do need to be careful about ascribing adjectives such as small, medium, and large to these effect sizes”; waarschuwt Hattie (2009, p. 9). Volgens Rubin (2010) zijn de Cohen's effectwaarden “approximate criteria, (...) merely rough guidelines” (p. 142). Cohen's classificatie moet dus niet al te rigide worden gehanteerd.²⁸¹ Desondanks stelt Magnusson (2014) dat het lijkt alsof wetenschappers die effectmaten klakkeloos overnemen. En ook Durlak (2009) stelt dat te veel auteurs de criteria exact, “iron-clad” (p. 922) toepassen. Terwijl Cohen zelf, nota bene, waarschuwde dat die maten ongeschikt zijn voor dergelijk algemeen gebruik (uit: Lipsey et al., 2012).

Volgens Baguley (2009) groeit de consensus dat dergelijke “canned” vuistregels (Lenth, 2001, p. 191) misleidend zijn bij het verklaren van effectgrootten en men ze daarom zou moeten vermijden. Een onderzoek van McCartney & Rosenthal (2000) liet bijvoorbeeld zien dat een effect-

grootte van amper .20, praktisch gezien, enorm statistisch significant kan zijn, terwijl dat effect *klein* is volgens een strikte hantering van Cohen's richtlijn (Lipsey et al., 2012). “Cooper, Charlton, Valentine en Muhlenbruck (2000) merken op dat een effectgrootte van 0.10, of lager zelfs, niet als onbelangrijk mag worden beschouwd” (uit: Marzano et al., 2012, p. 16). Zo schrijft Hattie (2009, p. 9) het volgende.

Further, there are many examples that show small effects may be important. A vivid example comes from medicine. Rosenthal and DiMatteo (2001) demonstrated that the effect size of taking low dose aspirin in preventing a heart attack was $d = .07$, indicating that less than one-eighth of one percent of the variance in heart attacks was accounted for by using aspirin. Although the effect size is small, this translates into the conclusion that 34 out of every 1,000 people would be saved from a heart attack if they used low dose aspirin on regular basis. This sounds worth it to me.

13.2.2 Praktische relevantie bij educatieve interventies

De creativiteitstrainingen die deel uitmaakten van de toetsingen bij dit *meervoudig onderzoek* zijn educatieve interventies. Het inzichtelijk maken van de *praktische relevantie* van het effect van dergelijke interventies, almede de interpretatie van effectgrootten, is eveneens niet vanzelfsprekend. Zo opteerde Wolf (1986) voor een *educatieve relevantie* bij een gestandaardiseerde gemiddelde effectgrootte van .25 en voor een *klinisch relevant effect* bij .05 (ook o.a.: Adnams et al., 2007; Neill, 2006). Rubin (2010) voegt daaraan toe: “Clinical significance [relevance] refers not only to the meaningfulness and practical value of the overall findings of a study – in terms of an intervention's *ES*, for example. It also refers to the meaningfulness and practical value of the benefits of an intervention for *each individual* recipient of the evaluated intervention” (p. 143). Ook Bloom, Hill, Bgraz & Lipsey (2008, pp. 6, 30), spreken van “educational significance [relevance]” bij een effectgrootte van .25. Volgens hen is Talmadge (1977, p. 34) de eerste die deze vuistregel noemt.

Daarnaast schrijft de *Koninklijke Nederlandse Akademie van Wetenschappen* (KNAW, 2009): “Een *ES* van .20 of hoger wordt door Slavin (2008) in onderwijsstudies als praktisch relevant gekwalificeerd” (p. 46). Rocconi & Gonyea (2015) noemen de “new reference values” (pp. 13-14), met de respectievelijke effectwaarden (*d*): $\geq .1$ is *klein*, $\geq .3$ *middelgroot* en $\geq .5$ *groot*. De *new reference values* zijn afgeleid uit een rapport van de *U.S. Department of Education* (Lipsey et al., 2012). Dat rapport is bedoeld voor onderzoekers van educatieve interventies, om hun statistische resultaten van effectgrootten en *praktische relevantie* begrijpelijk te kunnen presenteren aan beroepskrachten, beleidsmakers en eventueel andere onderzoekers die geïnteresseerd zijn in de geëvalueerde interventie. “Hattie (1999) wijst erop dat een effect moet worden afgezet tegen het gemiddelde effect van onderwijskundige interventies, en dat effect is 0,40” (uit: Driessen, 2007, p. 19). Hattie (2009) deed dat nogmaals in zijn boek *Visible Learning*, een synthese van ruim 800 meta-analyses, waar ongeveer 236 miljoen studen-

ten aan meewerkten. Relatief recentelijk schreef hij: “The typical growth effect-size per year is about 0.40, (...) with the usual cautions about context” (2015, p. 7).

Ellis (2010) stelt: “In the right context even small effects may be meaningful” (p. 35). Binnen onderwijsonderzoek bestaat de consensus dat kleine effectgrootten, althans volgens Cohen's vuistregels (zie *Voetmoot 88* en *Voetmoot 289*), toch tot zeer relevante verbeteringen kunnen leiden (o.a.: Shachar et al., 2003). Een sprekend voorbeeld daarvan is Coe's (2000b) uitspraak: “In education, if it could be shown that making a small and inexpensive change would raise academic achievement by an effect size of even as little as 0.1, then this could be a very significant improvement, particularly if the improvement applied uniformly to all students, and even more so if the effect were cumulative over time” (p. 5). Ellis (2010) noemt een tweede treffend voorbeeld: “Depending on what benefits can be achieved at what cost, an effect size of 2.0 might be ‘poor’ and one of .1 might be ‘good’ (Glass G.V. et al., 1981, p. 104)” (p. 41).

13.2.3 Interpretatie van effectgrootte & contextuele aspecten

Contextuele aspecten kunnen een belangrijke rol spelen bij de interpretatie van effectgrootten, wat dus van belang kan zijn voor de praktische relevantie. Neem bijvoorbeeld de kosten voor het implementeren van een nieuwe training in het onderwijs, waarvan uit onderzoek is gebleken dat het effect aan de lage kant is. Het denkbeeldige onderwijsinstituut zou desondanks de training graag willen continueren om de training te kunnen verbeteren. Deze wens maakt feitelijk het belang groter, dus ook de praktische relevantie. Bovendien blijken de kosten voor de implementatie van de training gering. Dat gegeven doet de praktische relevantie uiteraard extra toenemen. Omgekeerd zou het kunnen dat de implementatiekosten hoog zijn. Dat aspect is dan logisch van negatieve betekenis voor de praktische relevantie. Daar contextuele aspecten dus van belang kunnen zijn voor onderzoek van onderwijssituaties, licht ik dit onderwerp graag toe.

Zo komt, samengevat, uit de literatuur naar voren dat er geen algemene conventie is over het interpreteren en de relevantie van effectgrootten. Naast een combinatie van theoretische afwegingen, is het een kwestie van *common sense*, oftewel: *logisch verstand* (§ 8.2.1).²⁸² Binnen een bepaalde context bepalen onderzoekers de mate van relevantie van het effect van een interventie zelf, of het nu gaat om klinische, theoretische, biologische, educatieve of een andere vorm van *praktische relevantie* (o.a.: Kalinowski et al., 2010).

In de beschrijving van effectgrootte viel het woord *context* al een paar keer. Hierna wordt verder ingegaan op de contextafhankelijkheid van *praktische relevantie* en interpretatie van effectgrootte in relatie tot het huidige onderzoek. De eerste twee contextbepalende aspecten van dit onderzoek zijn het soort trainingen en het type interventies, respectievelijk: *creativiteitstrainingen* en *educatieve interventies*.

Deze sectie vervolg ik derhalve in drie thema's: *Kosten bij praktische relevantie* (§ A); *Andere relevante contextuele aspecten bij praktische relevantie* (§ B); en *Relevante contextuele aspecten uit het empirisch onderzoek* (§ C)

A: Kosten bij praktische relevantie

Een ander aspect is *Kosten*. “It need to be noted”, merkte Hattie (2009, p. 255) op, “that evidence based on effect sizes alone could lead to poor decisions. For any set of choices, there are costs as well as benefits. The financial costs of the various interventions may need to be taken into account when making decisions about what works best. It may be that we can use some of the cheaper interventions if their effects are positive, and this may be preferable to using some of the more expensive interventions.” Bij dit onderzoek gaat het met name om kosten die mogelijk gemoeid zijn met de implementatie van creativiteitstraining in het onderwijs, anders gezegd: in de curricula van *Bachelor of ICT*-opleidingen. Daarentegen zijn de belangen, dus ook de baten van creatievere BICT-studenten hoog (zie de algemene conclusie van *Deel I*).

Kosten is echter niet het enige criterium voor *educatieve significantie* (Tallmadge, 1977, pp. 33-34), ofwel: *educatieve relevantie*. Meerdere zaken kunnen daarop van invloed zijn. Dat blijkt onder meer uit Rubin's (2010) woorden:

Ultimately, substantive significance of any statistical finding depends on idiosyncratic value judgments about meaningfulness of the finding and its practical value to clients, significant others, society, or practitioners concerned about a problem to which that finding pertains. Those value judgments might consider any number of intangibles, such as whether the benefits of revising services or policies based on the implications of the finding outweigh the costs of the revisions, whether the variables assessed in the finding are really important, whether the finding adds much to what already known, and whether previous studies have other findings with more important implications” (p. 143).

Lipsey et al. (2012, p. 37), maken onderscheid tussen “cost-effectiveness” en “cost-benefit relationships”. De manier van representeren van het interventie-effect maakt het verschil. Een *kosten-/ effectiviteitsverhouding* vertegenwoordigt bijvoorbeeld het effect dat een interventie heeft op de leerprestatie van studenten. Wanneer de interventie te maken heeft met een direct geldelijk gewin, dan spreekt men van een *kosten-/batenverhouding*. Het gaat de reikwijdte van dit onderzoek te buiten, alsmede mijn expertise op bijvoorbeeld het gebied van accountancy, om gedetailleerde kostenanalyses te maken.²⁸³ Vandaar dat bij de interpretatie is uitgegaan van een algemene, *common sense* benadering (§ 8.2.1).

Hattie (2009, 2013, 2015) gaat bij educatieve interventies uit van een gemiddeld verwacht effect van het volgen van een jaar lang onderwijs. De tijdsduur van de huidige onderzochte creativiteitstrainingen staat in geen verhouding tot die periode van een jaar. Sommige trainingen nemen

slechts één dag *Deel In* beslag. Het spreekt voor zich dat de implementatiekosten navenant zijn en dus relatief laag. Daarnaast zijn deze creativiteitsstrainingen eenvoudig te implementeren in curricula, omdat professionele facilitators, als gastdocent van creativiteitstrainingen en/of de uitwerking van de lesmodules reeds beschikbaar zijn. Hetgeen betekent dat het onderwijsinstituut in kwestie geen ontwikkelingskosten heeft in dat opzicht en de kosten-baten-drempel als zodanig laag is; én dus direct daarmee aan de slag zou kunnen gaan.

B: Andere relevante contextuele aspecten bij praktische relevantie

Naast het gegeven dat de *implementatiekosten* van de hier bestudeerde *creativiteitstrainingen* (c.q. *educatieve creativiteit-bevorderende interventies*) betrekkelijk gering zijn, heeft dit onderzoek ook andere relevante contextuele aspecten laten zien die betekenisvol zijn bij de interpretatie van de effectomvang voor de *praktische relevantie* van het effect van creativiteitstraining bij BICT-studenten. Hier volgt een opsomming van die contextuele aspecten, in volgorde van de hoofdstukken en dus de deelvragen (zie het introducerend hoofdstuk) van dit proefschrift:

- Creativiteit is nodig voor een succesvol bestaan (*Hoofdstuk 1*). Creativiteit van mensen is in principe leerbaar (zie § 1.3.1 en *Hoofdstuk 3: Inleiding*);
- ICT is een *doorbraakinnovatie*, dus niet meer weg te denken uit de maatschappij. De ICT is innovatietopsectoren doorsnijndend en van groot belang voor onze economie. Zonder ideeën, en dus zonder creativiteit, is innoveren onmogelijk. Al tientallen jaren wordt bepleit dat creativiteit structureel in het onderwijs moet. Vanwege het sociaaleconomisch belang van ICT is creativiteitsontwikkeling in het bijzonder van belang voor ICT'ers, mede gezien het ontkoppingsprobleem (§ 6.4);
- Creativiteit & Innovatie is een pijler voor de hedendaagse (onderwijs-)vaardigheden (oftewel: de *21st Century Skills*); ICT en ICT-vaardigheden (of ICT-geletterdheid) zijn dat eveneens. Creativiteit ontbreekt bij de kernvakken (momenteel taal en rekenen), terwijl het cruciaal is bij Succesvolle intelligentie en *21st Century Skills*. Er is een groeiende behoefte aan ICT'ers. Creativiteitsontwikkeling ontbreekt in de Bachelor of ICT-domeinbeschrijving, waardoor het bevorderen van creativiteit niet is verankerd in BICT-curricula. Creativiteitsontwikkeling is in het bijzonder van belang voor Bachelors of ICT, vanwege het sociaaleconomisch belang van ICT en hun leidinggevende, creërende rol in de beroepspraktijk (zie *Hoofdstuk 4*);
- BICT-studenten lijken minder ideeën te bedenken dan studenten van een creatieve HBO-opleiding. Dat indiceert dat BICT-studenten minder creatief zijn en mogelijk een achterstand hebben wat betreft creativiteit dan HBO-studenten van een creatieve opleiding, anders gezegd: creativiteitstraining kan mede daardoor (extra) relevant zijn voor BICT-studenten (zie o.a. *Hoofdstuk 9*).²⁸⁴

282 (Zie tevens o.a.: Hattie, 2009; Kalinowski et al., 2010, p. 51; Slavin, 2008, p. 7)

283 Voor meer uitleg, zie het rapport van Lipsey et al. (2012, pp. 37-43), waarin tevens bronverwijzingen staan naar uitvoerige kostenberekeningen.

284 Zoals eerder aangegeven, betekent dit niet dat creativiteitstraining niet relevant kan zijn voor anderen (zie § 9.3: *Voetmoot 222*).

Samengevat: Het belang van creativiteitstraining voor BICT-studenten, oftewel het antwoord op de eerste onderzoeksvraag (§ 5.5), weegt in deze studie mee bij de interpretatie van de *praktische relevantie*.

C: Relevante contextuele aspecten uit het empirisch onderzoek

Ook bij de empirische studies van dit proefschrift zijn punten naar voren gekomen die van belang zijn voor de context van *praktische relevantie*. Ik noem er twee, respectievelijk *Fluencystudie* (§ C1); en *Enquêtestudie* (§ C2).

C1: Fluencystudie

De *Fluencystudie* toonde dat BICT-studenten over het algemeen minder ideeën bedenken bij *divergent denken-taken* (DDT's) dan studenten van creatieve HBO-opleidingen (zie § 9.2). Dat verschil viel vooral op bij de figuratieve DDT's. Dat leidt tot twee conclusies. Ten eerste, omdat BICT-studenten minder ideeën bedachten, zouden ze een achterstand kunnen hebben op dat vlak. Voor het wegwerken van die mogelijke achterstand kan creativiteitstraining relevant zijn voor hen. Ten tweede kan creativiteitstraining aannemelijk relevant zijn voor BICT-studenten, omdat visueel denkenden mogelijk creatiever zijn dan niet-visueel denkenden (§ 7.2.1), én BICT-studenten in dat opzicht aanzienlijk minder presteerden dan de HBO-studenten van creatieve opleidingen.

C2: Enquêtestudie

De relevantie van het volgen van creativiteitstraining kwam ook naar voren uit de beoordeling van de BICT-studenten bij de *Enquêtestudie*, gezien de vragenclusters: (a) “Belang”, (b) “Creatiever” en (c) “Intercollegiaal-inspirerend”.

Zo kreeg het vragencluster “Belang” (a) daarbij een totaal gemiddelde beoordeling van 4.21 ($Mdn = 5$) op de zespuntsschaal (§ 11.2.2). Daarnaast was bij de interventies met ervaren facilitators (ASP, CD en LD) de gemiddelde beoordeling van “Belang” $M = 4.63$, afgerond 5 op de zespuntsverdeling; en een effectgrootte $d_{unb} = 1.47$ (zie *Tabel 12* en *Tabel 17: pos.39*) in vergelijking tot de interventie met onervaren facilitators (CRT). Uit de gemeten cijfers van de respectievelijke enquêtevragen blijkt dat BICT-studenten er belang aan hechten creativiteitstraining te volgen voor zichzelf; maar ook dat het van belang is voor HBO-studenten van overige richtingen, én dat het mogelijk relevant is voor iedereen.

De andere twee vragenclusters (b) “Creatiever” ($d_{unb} = 1.08$) en (c) “Intercollegiaal-inspirerend” ($d_{unb} = 1.32$) - wederom bij de creativiteitstrainingen met ervaren facilitators - lijken eveneens factoren te zijn waarmee rekening gehouden moet worden binnen de context van het interpreteren van de effectomvang en de uiteindelijke bepaling van de *praktische relevantie*. (Voor de waarden, zie *Tabel 12* en *Tabel 17: pos.37 en 38*.)

Aldus houdt het cluster “Creatiever” (b) samengevat tevens in dat studenten vinden dat ze creatiever zijn geworden door de gevolgde creativiteitstraining. Daarnaast geeft het vragencluster “Intercollegiaal-inspire-

rend” (c) aan dat BICT-studenten vinden dat zij door creativiteitstraining beter zijn gaan samenwerken met collega's en dat ze meer inspiratie krijgen door elkaars ideeën, wat de collectieve, creatieve en innovatieve prestaties zou kunnen bevorderen. Dat kan in potentie gunstig zijn voor de studieprestaties en het functioneren in de (latere) beroepspraktijk. (Zie § 11.2.3 en § 11.3.4, voor meer informatie over *Clustering*.)

13.2.4 Sectiesamenvatting

In de vorige sectie (§ 13.1) is beschreven wat het belang is van *praktische relevantie* binnen dit onderzoek. Zo is bijvoorbeeld de *American Psychological Association* (APA, 2010) duidelijk over de noodzaak van de rapportage van effectomvang. In het algemeen representeert *praktische relevantie* de sterkte van een effect van *iets* op *iets*. In dit geval de effectomvang van creativiteitstraining bij *Bachelor of ICT*-studenten. Een eerste gevolgtrekking uit de inhoud van de paragraaf is dat er geen eenduidige effectmaten bestaan die bepalen of de uitwerking van een interventie bijvoorbeeld klein is, middelgoot, groot of zeer groot. De waarden uit bestaande vuistregels zijn dus geen vaste kantelpunten.

De omvang van effecten zijn contextafhankelijk. Daarom is bij de interpretatie van effectgrootten ten eerste gekozen om, naast Cohen's (o.a.: 1988, 1992) richtlijnen, rekening te houden met praktische regels van anderen, zoals die van Wolf (1986) en Bloom et al. (2008), die een Cohen's d van .25 classificeren als *educatief relevant*. Slavin (2008) doet dat bij .20. De *new reference values* van Lipsey et al. (2012) zijn: $d \geq .1$ klein, $\geq .3$ middelgroot en $\geq .5$ groot. En Hattie (o.a.: 2009), als laatste voorbeeld, kwalificeert een gemiddelde effectmaat van .40 als educatief relevant. Hattie's waarde is mede interessant, omdat het als uitkomst verwacht mag worden van het volgen van een jaar lang onderwijs.

Daarnaast moeten de genoemde bevindingen (§ 13.2.1) van Coe (2000b), Ellis (2010)²⁸⁵ en Marzano & Waters (2012)²⁸⁶ serieus worden genomen. Zij gaven namelijk aan dat een effectgrootte van 0.10, of zelfs lager, relevant kan zijn; zeker in situaties waarbij de opbrengsten relatief hoog zijn en de kosten laag.

Bovendien spelen contextuele aspecten, waarvan kosten en baten deel uitmaken, een relevante rol bij de interpretatie van de *praktische relevantie* en de uiteindelijke effectomvang. Contextbepalend bij dit onderzoek is onder meer dat de onderzochte interventies *creativiteitstrainingen* zijn en dat het gaat om *educatieve interventies*. Daarnaast betreft de context samenvattend: persoonlijke, didactische en sociaaleconomische aspecten (zie bovenstaand overzicht) én de genoemde aspecten die uit het empirisch onderzoek naar voren zijn gekomen die relevant zijn bij de respectievelijke *kosten-effectiviteit-* en *kosten-/batenverhouding*.

Het is een gegeven dat zowel de ontwikkelings- als de implementatiekosten relatief gering zijn van de empirisch onderzochte creativiteitstrainingen die

deel uitmaakten van dit meervoudige onderzoek. Ten eerste, omdat de facilitators die de trainingen verzorgden de betreffende training reeds hebben ontwikkeld. Ten tweede, omdat zij in te huren zijn als gastdocent tegen relatief lage kosten. En ten derde, omdat de onderzochte creativiteitstrainingen uiteenlopen van slechts één tot tien dagdelen, terwijl de effectsterkte van educatieve interventies veelal gemeten wordt na een jaar lang onderwijs (zie o.a.: Hattie, 2009).

Kortom, omdat de kosten relatief laag zijn, én de mogelijke opbrengsten (waaronder de economische) daarentegen relatief hoog, is bij de interpretatie van de effectomvang en het bepalen van de *praktische relevantie* bovendien rekening gehouden met de mogelijke relevantie van effectgrootten van .10, of eventueel lager (zie § 13.2.1). “Small effects can be important if they trigger big consequences” (Ellis, 2010, p. 35). Bijvoorbeeld de economisch groei die zou kunnen ontstaan door het toenemend aantal innovaties van creatievere *Bachelors of ICT*, als effect van de door hen gevolgde creativiteitstrainingen.

Tot slot van deze sectie (§ 3.2) moet gezegd dat de resultaten van eerdergenoemde meta-analytische studies van derden (*Hoofdstuk 3: Inleiding*) een context vormen waar de effectwaarden van het huidige onderzoek mee vergeleken zullen worden. Al met al vertegenwoordigen de genoemde gezamenlijke contextuele aspecten (waaronder dus de essentie van creativiteitstraining voor BICT-studenten), in deze studie, een relatief groot belang bij het interpreteren van effectgrootten en het bepalen van *praktische relevantie*.

“The evolution of civilization depends on innovation, and innovation depends on creativity. In the economy, innovation is decisive for one product to gain a share of the market. Therefore, it is undoubtedly important to nurture and enhance creativity in students”

Hsen-Hsing Ma (2006, p. 435), Department of Education, National Chengchi University.

13.3 Gecombineerde effectomvang & praktische relevantie

Al eerder is toegelicht dat gezond verstand (*common sense*) de meest eenvoudige manier is om meetresultaten te interpreteren (zie o.a. § 8.2.1). Daarbij is ook gezegd dat gezond verstand alleen vaak niet voldoende is; en dat er veelal andere middelen gebruikt worden om het gezond verstand te onderbouwen (zie § 8.2.2).

Daarom beschrijf ik in deze sectie de samenvoeging (onderzoekssynthese) van de resultaten uit de drie studies van dit meervoudige empirisch onderzoek (zie § 8.4, voor de introductie daarvan). Die drie studies zijn uitgevoerd om antwoorden te vinden op de tweede onderzoeksvraag: “Werkt creativiteitstraining die studenten *Informatie- en Communicatietechnologie* van Nederlandse hogescholen volgen vanuit het curriculum?”

De onderzoekssynthese is tot stand gekomen op grond van de effectgrootten (effect size: ES, Cohen's d) uit de Interventiestudie (*Hoofdstuk 10*), de Enquêtestudie (*Hoofdstuk 11*) én de Expertsstudie (*Hoofdstuk 12*). Ook zijn de p-waarden uit die studies overgenomen. Om aanvullend inzicht te bieden in de gezamenlijke effecten, zijn tevens gemiddelde effectgrootten berekend (Cunningham's d_{unb}). Daarvoor is “simple mean effect size” (Ellis, 2010, p. 102) gebruikt. Middels de synthese ontstaat er een totaalbeeld vanuit een kwalitatief perspectief. De resultaten zijn samengevat in *Tabel 17*.

285 (C.q.: Glass G.V. et al., 1981)

286 (C.q.: Cooper H.M. et al., 2000)

Tabel 17. De onderzoekssynthese van het meervoudig empirisch onderzoek naar antwoorden op de tweede onderzoeksvraag: "Werkt creativiteitstraining die studenten Informatie- en Communicatietechnologie van Nederlandse hogescholen volgen vanuit het curriculum?"

Pos	Studie	Aanpak	Design	Criterium	Training	DD-taak / Vraagcluster	ES	C.I. 95% (d _{sub})	Sig	
		DDT	prepost		ASP**	tekstueel figuratief samen	d _{sub}	min	max	
01	Interventiestudie			Aantal verschillende ideeën (fluency), Leerwinst (post - pre)	ASP**	tekstueel figuratief samen	.57	-.19	1.32	.002
02							.33	-.41	1.08	.004
03							.50	-.25	1.26	<.001
04				Idem (Controlegroep-a)	Ctrl-a	tekstueel figuratief samen	.06	-.68	.80	.730
05							.28	-.46	1.02	.195
06				Idem (ASP - Controlegroep-a), Netto leerwinst*		tekstueel figuratief samen	.23	-.52	.97	.211
07							.51			.032
08						figuratief samen (a)	.05			.400
09							.28			.063
10				Aantal verschillende ideeën (fluency), Experimentele vs Controlegroep-b	CD**	tekstueel figuratief samen (b)	.69	.06	1.33	.031
11							.41	-.22	1.04	.195
12							.63	-.01	1.26	.051
13					CRT***	tekstueel figuratief samen (c)	.32	-.31	.94	.313
14							-.12	-.74	.50	.703
15							.09	-.53	.71	.767
16				Aantal verschillende ideeën (fluency), Experimentele groepen (CD + CRT) vs Controlegroep-b		tekstueel figuratief samen	.50	-.05	1.04	.071
17							.16	-.38	.70	.554
18							.37	-.17	.92	.171
19				*Netto gemiddelde effectomvang van de prepost- en postonly-Interventie-studie (ofwel: totale Interventiestudie)		tekstueel figuratief samen	.50			
20							.11			
21							.33			
22				*Idem, enkel ervaren facilitators** (das ASP en CD)		tekstueel figuratief samen	.60			
23							.23			
24							.45			
25	Enquêtestudie** (Ervaren vs Onervaren)	CSA	enquête	Studenten-beoordeling m.b.t. vraagclusters** (ASP** vs CRT***)	ASP**	*Creatiever" "Intercollegaal/Inspireerend" "Behang" Clusters gezamenlijk (a)	1.10	.48	1.71	<.001
26							1.52	.87	2.17	<.001
27							1.51	.87	2.16	<.001
28							1.40	.76	2.03	<.001
29				Idem** (CD** vs CRT***)	CD**	*Creatiever" "Intercollegaal/Inspireerend" "Behang" Clusters gezamenlijk (b)	.73	.05	1.41	.034
30							1.22	.51	1.94	.001
31							1.31	.59	2.03	<.001
32							.97	.28	1.67	.005
33				Idem** (LD** vs CRT***)	LD**	*Creatiever" "Intercollegaal/Inspireerend" "Behang" Clusters gezamenlijk (d)	1.30	.55	2.05	.001
34							1.25	.50	1.99	.001
35							1.37	.62	2.13	<.001
36							1.41	.65	2.18	<.001
37				Gemiddelde effectomvang Enquêtestudie** (ASP, CD en LD samen** vs CRT***)	Samen	*Creatiever" "Intercollegaal/Inspireerend" "Behang" Clusters gezamenlijk	1.08	.54	1.61	<.001
38							1.32	.77	1.87	<.001
39							1.47	.91	2.03	<.001
40							1.31	.76	1.86	<.001

41	Experts-studie	CAT	prepost	Creativiteitsbeoordeling (door experts), Leerwinst (pre - post)	ASP**	tekstueel figuratief samen	.16	-.58	.91	.470
42							.48	-.27	1.23	.007
43							.38	-.37	1.13	.064
44				Idem (Controlegroep-a)	Ctrl-a	tekstueel figuratief samen	-.41	-1.15	.34	.090
45							.45	-.30	1.20	.080
46							-.02	-.76	.72	.905
47				Idem (ASP vs Controlegroep-a), Netto leerwinst*		tekstueel figuratief samen (a)	.57			.066
48							.03			.882
49							.40			.178
50				Creativiteitsbeoordeling (door experts), Experimentele groepen (CD + CRT) vs Controlegroep-b	CD**	tekstueel figuratief samen (b)	-.14	-.76	.48	.693
51							.66	.02	1.29	.041
52							.33	-.29	.95	.263
53					CRT***	tekstueel figuratief samen (c)	.28	-.34	.91	.357
54							-.17	-.79	.45	.587
55							.07	-.55	.69	.818
56				Creativiteitsbeoordeling (door experts), Experimentele groepen (CD + CRT) vs Controlegroep-b		tekstueel figuratief samen	.07	-.47	.61	.750
57							.26	-.28	.79	.360
58							.20	-.34	.74	.444
59				*Netto gemiddelde effectomvang van de prepost- en postonly-Experts-studie (ofwel: totale Experts-studie)		tekstueel figuratief samen	.32			
60							.14			
61				*Idem, enkel ervaren facilitators (das ASP en CD)**		tekstueel figuratief samen	.22			
62							.35			
63							.36			
64							.69	.34		
65	Synthese (Totaalresultaat)*	triangulatie		*Sub gemiddelde effectomvang, triangulatie per creativiteitstraining (samen)	ASP** CD** CRT*** LD**	(a) triangulatie samen (b) idem (c) enkel Interventie- & Experts-studie (d) uitsluitend bij de Enquêtestudie	.69 .64 .08 1.41			
66							.31			
67							.41			
68				*Sub gemiddelde effectomvang incl. CRT enkel Interventie- & Experts-studie			.66			
69				*Idem (ervaren facilitators: ASP en CD samen)**			.91			
70				*Idem, nu met "Creatiever" uit de Enquêtestudie**			.67			
71							.66			
72				*Totaal gemiddelde effectomvang (ervaren facilitators)**			.91			
73				* Idem, echter zonder LD i.v.m. uitsluitend bij de Enquêtestudie			.67			
74				* Totaal gemiddelde effectomvang (synthese van alle studies tezamen)			.71			
75				* Idem, echter weer zonder LD i.v.m. uitsluitend bij de Enquêtestudie			.47			

Note. ES = effect size (effectgroote); *Cumming's d unbiased (d_{sub})*; Sig = Statistische significantie (p-waarde), overgenomen uit de lopende studies. * = Gecombineerde gemiddelde effectgroote (simple mean effect size). **Interventie met ervaren facilitators = ASP, CD en LD (enkel bij de Enquêtestudie heeft CRT gefungeerd als controlegroep); ***Interventie met onervaren facilitators = CRT. (e) = enkel Interventie- & Experts-studie. Aanpak: DDT = Divergent Denken-Test; CSA = Creative Self Assessment; en CAT = Consensual Assessment Technique.

Om meer inzicht te bieden in de resultaten van de *Praktische-relevantie-studie* vervolg ik deze sectie met de beschrijving van drie thema's: *Gecombineerde effectgrootten* (§ 13.3.1); *Praktische relevantie van de effectwaarden* (§ 13.3.2); en *Gemiddelde effectomvang per creativiteitstraining* (§ 13.3.3).

13.3.1 Gecombineerde effectgrootten

De onderzoekssynthese van de resultaten van het meervoudig empirisch onderzoek laat over het algemeen zien dat creativiteitstraining een positief effect heeft – dus dat het werkt - bij BICT-studenten, én dat de *praktische relevantie* sterk is binnen de context. Zo wijzen de resultaten van zowel de *Interventiestudie*, de *Enquêtestudie*, als de *Experts-studie* in die richting. Om een zo nauwkeurig mogelijk beeld te schetsen, volgt een gespecificeerde uiteenzetting, mede bedoeld als toelichting op *Tabel 17*.

De totale algemene gemiddelde effectomvang $d_{\text{unb}} = .71$ (zie *pos.74* in *Tabel 17*) is een indicatie van dat algemeen positieve effect. Er moet direct worden aangetekend dat die waarde vertekend is, omdat de creativiteitstraining *Lenig Denken* (LD) enkel is onderzocht bij de *Enquêtestudie*, met een effectgrootte van $d_{\text{unb}} = 1.41$ (*pos.36*). Zonder LD is die effectgrootte $d_{\text{unb}} = .47$ (*pos.75*). Het dus niet geheel realistisch de eerstgenoemde uitkomst, $d_{\text{unb}} = .71$, als geharnast gemiddeld effect te beschouwen. Een tussenliggende waarde is aannemelijker.

De gecombineerde gemiddelde effectgrootte $d_{\text{unb}} = .91$ (*pos.72*) van creativiteitstrainingen mét *ervaren facilitators* veronderstelt een meer betrouwbare waarde dan de hierboven genoemde $d_{\text{unb}} = .71$ van *pos.74*.²⁸⁷ Toch is dat niet geheel zo, omdat ook de effectgrootte van LD daarin is verdisconteerd terwijl die training enkel deel uitgemaakte van de *Enquêtestudie*. Hetgeen bias geeft en dus een scheef beeld. Bovendien diende *Creatieve Technieken* (CRT) als controlegroep bij die studie, waardoor het effect van CRT buiten beschouwing is gebleven. Wanneer het effect van LD wordt weggelaten uit dat resultaat dan geeft dat een gemiddelde effectwaarde van $d_{\text{unb}} = .67$ (*pos.73*). In dat geval weegt het resultaat van de *Enquêtestudie* dus niet mee, wat eveneens bias geeft.

Een mogelijk realistischer beeld ontstaat wanneer de uitkomsten van de *Enquêtestudie* deels worden verdisconteerd in de totale gemiddelde effectomvang. Bij de berekening daarvan is enkel uitgegaan van de effectgrootten van het vragencluster “Creatiever” (*pos.25* en *pos.29*), aangezien dit cluster uiteindelijk binnen de vraagstelling valt of BICT-studenten creatiever geworden zijn door creativiteitstraining. Daarentegen heeft cluster “Intercollegiaal-inspirerend” een sociale context; en gaat het vragencluster “Belang” over de relevantie van creativiteitstraining. Uiteindelijk leverde de berekening (dus ook zonder CRT en LD) een gemiddelde gecombineerde effectomvang op van $d_{\text{unb}} = .66$ (*pos.71*).

Daarnaast heb ik ook de totale gemiddelde effectwaarde zonder inbreng van het cluster “Creatiever” in het overzicht vermeld, $d_{\text{unb}} = .31$ (*pos.69*). Anders gezegd: daarin hebben enkel de effectgrootten van de *Interventiestudie* en de *Experts-studie* meegewogen.

13.3.2 Praktische relevantie van de effectwaarden

De gemiddelde gecombineerde effectomvang van $d_{\text{unb}} = .66$ (*pos.71*) is conform een range van meta-analytische bevindingen van derden (zie *Hoofdstuk 3: Inleiding*). Zo gaf de studie van Scott, Lertz & Mumford (2004a) een gezamenlijke gemiddelde effectgrootte van $d = .68$. Een andere meta-analyse van hen (Scott et al., 2004b) liet $d = .77$ zien als gemiddelde effectwaarde. De totale gemiddelde effectsterkte van creativiteitstraining bij Ma's (2006) metastudie was $d = .77$. Tot slot noteert Hattie (2009) $d = .65$ als totaal gemiddeld effect van creativiteitsprogramma's na een jaar lang onderwijs. Uitgaande van Cohen's vuistregels vallen deze effectwaarden in het gebied van middelgroot ($d = .50$) tot groot ($d = .80$).²⁸⁸

Bij hantering van de richtlijnen voor educatieve interventies, die per auteur dus variëren van $d = .20$ tot $d = .40$ (§ 13.2.2), is de effectsterkte van $d_{\text{unb}} = .66$ (*pos.71*) relatief groot te noemen. Wanneer tevens de lage kosten én de in potentie grote, persoonlijke, didactische en sociaaleconomische baten in de beoordeling worden meegewogen, dan is er bovendien sprake van een relatief zeer omvangrijke *praktische relevantie*. Zeker gezien de relatief korte tijdspanne van de in dit proefschrift onderzochte creativiteitstrainingen - variërend van één tot een tiental dagdelen - in relatie tot een jaar lang onderwijs, waarvan Hattie is uitgegaan bij het bepalen van *praktische relevantie* bij zijn onderzoek naar educatieve interventies.

Zelfs wanneer de resultaten van de *Enquêtestudie* buiten beschouwing worden gelaten, dan is gemiddelde effectgrootte $d_{\text{unb}} = .31$ (*pos.69*) nog altijd *praktisch relevant* in educatief opzicht. Die relevantie geldt in het bijzonder als er daarnaast rekening wordt gehouden met de persoonlijke, didactische en sociaaleconomische baten tegenover de relatief geringe kosten; én het gegeven dat de tijdsduur van de onderzochte trainingen in verhouding een fractie is van de tijdspanne die Hattie hanteerde bij zijn studie. De totale implementatiekosten van de hier onderzochte creativiteitstrainingen zijn dus relatief laag.

13.3.3 Gemiddelde effectomvang per creativiteitstraining

Om het inzicht in het effect van creativiteitstraining te vergroten, volgt hier tevens een beschrijving van de resultatensynthese per onderzochte creativiteitstraining: *Aware Sensory Perception* (ASP), *Creatief Denken* (CD), *Lenig Denken* (LD) en *Creatieve Technieken* (CRT).²⁸⁹ Daarnaast wordt

het effect belicht van de categorie waarbinnen deze trainingen vallen (zie § 5.2.5).

Een eerste blik op de effectgrootten van de verschillende creativiteitstrainingen laat zien dat er grote verschillen bestaan tussen de creativiteitstrainingen (*pos.65-68*). Die getallen geven echter een vertekend beeld omdat niet elke interventie deel uitmaakte van hetzelfde experiment, waardoor er dus verschillende uitkomsten zijn gebruikt voor het bepalen van de respectievelijke effectgrootten. Voor het bepalen van vergelijkbare effectgrootten van ASP, CD en CRT zijn enkel de gezamenlijke resultaten van de *Interventie-* en de *Experts-studie* beschouwd. Voor ASP is de gecombineerde gemiddelde effectwaarde $d_{\text{unb}} = .34$ (*pos.65e*); voor CD $d_{\text{unb}} = .48$ (*pos.66e*) en voor CRT $d_{\text{unb}} = .08$ (*pos.67*). Opvallend is het relatief grote verschil tussen ASP en CD enerzijds en CRT anderzijds. LD wordt hier niet genoemd, omdat die training enkel deel uitmaakte van de *Enquêtestudie*.²⁹⁰ Daarnaast geldt dat CRT de controlegroep was bij de *Enquêtestudie* en daarin dus geen effectgrootte kent.

Op meerdere plaatsen in dit proefschrift is aangehaald dat de onervarenheid van de CRT-studenten om op te treden als facilitator een reden zou kunnen zijn voor de lagere bevindingen vergeleken met die van de andere creativiteitstrainingen. Zo kan ook hier het relatief grote verschil tussen ASP en CD aan de ene kant en CRT aan de andere kant een aanwijzing zijn in die genoemde richting. Ter indicatie, van wat het effect kan zijn van een creativiteitstraining die in handen is van ervaren facilitators, is ook die effectgrootte berekend $d_{\text{unb}} = .41$ (*pos.70*).²⁹¹ Daarbij wogen dus enkel de resultaten mee van de *Interventie-* en de *Expertstudie*, CRT uitgezonderd.

Hierna ga ik gedetailleerd in op de gemiddelde effectomvang van de drie creativiteitstrainingen (categorieën): *Aware Sensory Perception*, categorie: *Attitude Training* (§ A); *Creatief Denken & Lenig Denken*, categorie: *creative problem solving* (§ B); en *Creatieve Technieken*, categorie: *Studiemodule Facilitator* (§ C).

A: *Aware Sensory Perception* (*Attitude Training*)

Het totale gecombineerde gemiddelde effect van de creativiteitstraining *Aware Sensory Perception* (ASP) is $d_{\text{unb}} = .69$ (*pos.65*).²⁹² Wanneer enkel de mening over creativiteitsbevordering van BICT-studenten door ASP zou gelden, of anders gezegd: alleen het vragencluster “Creatiever” zou meewegen bij de controle van de gecombineerde effectgrootte, dan is de totale effectomvang $d_{\text{unb}} = .61$ (het gemiddelde van *pos.09*, *25* en *49*). Beide effectwaarden betekenen een hoge praktische significantie, uitgaande van educatieve vuistregels en de overige context (oftewel: de persoonlijke, di-

didactische en sociaaleconomische baten in verhouding tot de relatief lage kosten). Daarnaast geldt voor ASP dat de tijdsduur van tien dagdelen (oftewel de vijf werkdagen die de module in beslag neemt) zich niet verhoudt tot Hattie's vuistregel die uitgaat van een te verwachten effectgrootte van .4 bij een jaar lang onderwijs. Ook wanneer dat meeweegt bij de interpretatie is er sprake van een relatief zeer hoge *praktische, educatieve relevantie*, waardoor de creativiteitstraining ASP onverlet in BICT-curricula geïmplementeerd zou kunnen worden.

ASP behoort tot de categorie van creativiteitstrainingen *Attitude Training* (§ 5.2.5). ASP was de enige onderzochte training uit deze categorie. Daarom gelden de voorgaande bevindingen over *Aware Sensory Perception* ook voor het type *Attitude Training* binnen dit onderzoek. Samengevat wijzen de resultaten van ASP in de richting van een positief effect bij BICT-studenten.

B: *Creatief Denken & Lenig Denken* (*creative problem solving*)

De trainingen *Creatief Denken* (CD) en *Lenig Denken* (LD) vallen in de categorie *creative problem solving* (CPS), zie § 5.2.5. Daarom staan beide trainingen hierboven onder één noemer. De gemiddelde gezamenlijke CPS-effectomvang bij deze studie is $d_{\text{unb}} = .83$, het gecombineerde gemiddelde van LD (*pos.36* of *pos.68*) en CD (*pos.12*, *32* en *52*). Dit is vertekend omdat LD enkel deel uitgemaakte van de *Enquêtestudie*, al blijft LD natuurlijk behoren tot de categorie CPS. Wanneer de effectgrootte van LD desondanks buiten beschouwing blijft, dan is de gecombineerde gemiddelde effectomvang van CPS (dus feitelijk enkel van CD) $d_{\text{unb}} = .64$ (*pos.66e*).²⁹³ Als wederom enkel naar de werking van de training bij BICT-studenten gekeken wordt qua creativiteit, dan is de totale gemiddelde effectsterkte van CD $d_{\text{unb}} = .56$ (het gemiddelde van *pos.12*, *29* en *52*). Oftewel: het effect is gelijk aan de vorige waarde, maar dan weer zonder de vragenclusters “Intercollegiaal-inspirerend” en “Belang”. Deze gemiddelde effectwaarden van CD betekenen dat er – wederom uitgaande van educatieve vuistregels en de persoonlijke, didactische en sociaaleconomische baten in verhouding tot de relatief lage kosten - sprake is van een hoge praktische relevantie.

De *praktische, educatieve relevantie*, is relatief zeer hoog te noemen, wanneer daarin wordt verdisconteerd dat CD bij dit onderzoek slechts één dagdeel duurde, weer in tegenstelling tot de vuistregel van Hattie die uitgaat van .4 als verwachte gemiddelde effectgrootte van een jaar lang onderwijs. Daaruit mag worden geconcludeerd dat *Creatief Denken* (CD) zich mogelijk leent om in BICT-curricula te worden geïmplementeerd. Dat geldt mogelijk ook voor *Lenig Denken* (LD), gelet op de mening van de studenten over LD (zie *Hoofdstuk 11* en *pos.36* van de tabel); en wellicht over het algemeen voor andere trainingen van het type *creative problem solving* (CPS).

287 Voor meer toelichting op onervaren en ervaren facilitators, zie o.a. § 5.2.5, § 10.3.2, § 11.3.2 en § 11.4.

288 Volgens de consensus-vuistregel van Cohen (o.a.: 1988, 1992), is er bij een (Cohen's) d van ongeveer .20 sprake van een klein effect; vanaf .50 van een medium effect; en bij .80 van een groot effect. Deze vuistregel is een globale richtlijn en moet niet exact (geharnast) worden toegepast. Voor meer informatie over effectgrootte en de interpretatie daarvan, zie o.a.: *Voetnoot 86*, § 8.2 en § 13.2.

289 Zie § 5.2.1 - § 5.2.5 voor meer toelichting op de inhoud van de onderzochte creativiteitstrainingen; en idem dito § 3.1 op de soorten creativiteitstraining.

290 Alhoewel de opzichzelfstaande resultaten bij de *Enquêtestudie* (§ 11.2.3) tonen dat er wel degelijk ook een groot verschil is tussen LD en CRT.

291 Voor de duidelijkheid: het afgerond gemiddelde van *pos.24* en *pos.64*.

292 De effectgrootte van de *Enquêtestudie* is daarin verdisconteerd.

293 Ook nu is de effectgrootte van de *Enquêtestudie* erin verdisconteerd.

C: Creatieve Technieken (Studiemodule Facilitator)

Ondanks dat *Creatieve Technieken* (CRT) uitgaat van CPS-principes behoort deze training feitelijk tot de categorie creativiteitstrainingen *Studiemodule Facilitator* (§ 5.2.5). CRT is namelijk een studiemodule, waarbij aan studenten de grondbeginselen van het vak *facilitator* van CPS worden bijgebracht (zie tevens § 5.2.2). De studenten moesten als eindopdracht zelf een creatieve sessie faciliteren. Bij dit onderzoek is het effect gemeten van die eindopdracht (de interventie). Er kan van studenten niet worden verwacht dat zij een creatieve sessie net zo doeltreffend leiden als ervaren facilitators. Dat lijkt ook naar voren te komen uit de resultaten van het meervoudig empirisch onderzoek. De drie losse studies vertonen namelijk een dergelijke overeenkomstige trend. Te weten dat CRT (relatief veel) lager scoorde dan de andere creativiteitstrainingen. Dat heb ik geïnterpreteerd als een indicatie dat de *onervarenheid* van de studenten, om een interventie met identiek effect te leiden als professionele (ervaren) facilitators, een mogelijke oorzaak zou kunnen zijn van die trend (zie o.a. § 10.4, § 11.4 en § 12.5).

De totale algemene gemiddelde effectomvang van CRT is $d_{\text{unb}} = .08$ (pos.67), mede door het niet kunnen controleren van een effectgrootte bij de *Enquêtestudie*.²⁹⁴ Hetgeen niet zegt dat CRT helemaal geen effect had bij de studenten. CRT werkte namelijk wel voor hen bij het bedenken van ideeën bij TDDT: $d_{\text{unb}} = .32$ (pos.13) bij de *Interventiestudie* en $d_{\text{unb}} = .28$ (pos.53) bij de *Experts-studie*, anders gezegd: met $d_{\text{unb}} = .30$ als gemiddeld effect. De relatief lage gemiddelde effectomvang is veroorzaakt door de negatieve totale gemiddelde effectgrootte van $d_{\text{unb}} = -.15$ bij de figuratieve DDT's; respectievelijk: $d_{\text{unb}} = -.12$ (pos.14) bij de *Interventiestudie* en $d_{\text{unb}} = -.17$ (pos.54) bij de *Experts-studie*. Daarnaast kan daaruit worden opgemaakt dat de studenten die CRT volgden mogelijk minder creatief zijn dan de andere studenten, omdat minder creatieven doorgaans minder visueel - dus mogelijk minder beeldend (minder figuratief denkend) - zijn ingesteld (zie § 7.2.1). Samengevat had CRT positief effect bij BICT-studenten, al is dat slechts op het bedenken van tekstuele ideeën.

Volgens educatieve vuistregels zijn de gemiddelde effectwaarde en de effectwaarden bij de TDDT's *praktisch relevant*. Enerzijds mede omdat CRT ontwikkeld is vanuit de vraag van het beroepenveld naar ICT'ers die creatieve sessies kunnen faciliteren (§ 5.1). Anderzijds omdat de persoonlijke, didactische en sociaaleconomische baten en de lage kosten van positieve betekenis zijn voor de praktische relevantie van CRT. Voorts moet rekening gehouden worden met het gegeven dat CRT een studiemodule is met een looptijd van *slechts één* lesperiode (maximaal 10 weken, à ongeveer twee uur = een half dagdeel), in plaats van een heel schooljaar, waarvan

Hattie, zoals eerder toegelicht, uitgaat bij de interpretatie van de effectomvang van educatieve interventies en bij het bepalen van de praktische, educatieve relevantie. Bij dit alles dient erop te worden gewezen dat ook kleine effectgrootten, bijvoorbeeld $d_{\text{unb}} = .10$ of zelfs lager, praktisch relevant kunnen zijn, in het bijzonder wanneer de opbrengsten hoog zijn en de kosten laag, zoals bij CRT.

Op grond van de praktische, educatieve relevantie, kunnen beleidsmakers, curriculumontwikkelaars en andere belanghebbenden in overweging nemen om de studiemodule *Creatieve Technieken* (CRT) te implementeren in *Bachelor of ICT*-curricula, niet in de laatste plaats omdat CRT inspeelt op behoeften uit het ICT-beroepenveld. Desondanks luidt het advies CRT verder te onderzoeken en indien mogelijk de module te verbeteren.

13.4 Bevindingen & discussies

Gefundeerd op de resultaten van de resultatensynthese van de *Praktische-relevantiestudie* - waarvan de totale uitkomst in overeenstemming is met de veronderstellingen van de deelstudies (zie *Hoofdstuk 10, 11 & 12*); oftewel: de hypothesen van die studies niet verworpen konden worden - luidt de algemene conclusie dat creativiteitstraining werkt bij studenten *Informatie- en Communicatietechnologie* van Nederlandse hogescholen (BICT-studenten), mede gezien de uitkomsten van meta-analytisch onderzoek van anderen.

Dat geldt in het bijzonder voor creativiteitstrainingen waarvan de interventie in handen was van ervaren, professionele facilitators: *Aware Sensory Perception* (ASP), *Creatief Denken* (CD) en *Lenig Denken* (LD). Het effect van *Creatieve Technieken* (CRT), de training die door de studenten zelf is verzorgd (onervaren facilitators), was gemiddeld relatief klein ($d_{\text{unb}} = .08$, pos.67). Al moet daarbij worden opgemerkt dat de mening van de studenten niet in die effectomvang heeft meegewogen, omdat CRT bij de *Enquêtestudie* als controlegroep fungeerde.

Daarnaast heb ik betoogd dat educatieve interventies met een kleine effectsterkte toch *praktisch relevant* kunnen zijn. Zo is in de voorgaande paragraaf beschreven dat dit zelfs geldt voor CRT, gezien de context: (i) de vraag naar CRT uit het ICT-beroepenveld; (ii) de persoonlijke, didactische en sociaaleconomische voordelen; en (iii) de lage implementatie- en ontwikkelingskosten.

Al met al wijzen de resultaten van de separate studies in de richting van een positief effect voor creativiteitstraining bij BICT-studenten. Echter bij CRT is het een gegeven dat dit uitsluitend geldt voor de gemeten resultaten van de TDDT's. Anders geformuleerd: Ik weet niet waarom, maar de gemeten interventie (de eindtoets van CRT) liet een nihil tot geen effect

zien bij de deelnemende BICT-studenten. Desondanks vind ik dat er steun is voor de onervarenheid van de studenten die de interventie leidden, als oorzaak van dat geringe effect.

13.4.1 Discussie over de toelichting op statische significantie & praktische relevantie

Ergens in het traject van dit proefschriftonderzoek viel mij op dat er meer is dan enkel een *p-waarde* (*statistische significantie*) bij statistisch onderzoek. Namelijk *effectsterkte* (*effect size*, *Es*), een maat voor *praktische relevantie* dat - volgens de literatuur geheel ten onrechte - veelal ook *praktische significantie* wordt genoemd (zie § 13.1). Bovendien viel mij op dat er een wezenlijk verschil bestaat tussen *statistische significantie* en *praktische relevantie*. Vandaar dat ik het verschil heb toegelicht.

Toch wil ik hier graag benadrukken dat die toelichting niet is bedoeld om een discussie te beginnen op statistisch-theoretisch terrein, daar dat een onderwerp is voor statistisch theoretici. Dat valt buiten mijn bereik en buiten de doelstelling van dit proefschriftonderzoek. Echter, door mijn constatering dat het woord *significant* ambigu is in onderzoek, en bovendien verwarring in de hand werkt, vond ik verslaggeving ervan onontbeerlijk. Zeker, omdat *praktische relevantie* - anders gezegd het effect van iets op iets - essentieel is voor dit onderzoek, in dit specifieke geval de werking van creativiteitstraining bij BICT-studenten.

13.4.2 Verdere discussies

Vooropgesteld geldt dat de meeste discussies omtrent deze resultatensynthese geïntegreerd zijn gevoerd in de voorgaande paragrafen en in de conclusiesecties van de drie deelstudies in kwestie (zie *Hoofdstuk 10, 11 en 12*), uitgezonderd de discussie over de gemiddelde effectgrootten uit dit hoofdstuk. Daarvoor is, zoals gezegd, *simple mean effect size* gebruikt (§ 13.3). Daar is voor gekozen, ondanks dat het vertekening (bias) in de resultaten teweeg kan brengen.

De steekproefomvang laten meewegen bij de berekening is onder meer een manier om die bias te verkleinen. Men spreekt dan van "*weighted mean* [of *average*] *effect size*" (o.a.: Ellis, 2010, p. 103). Middels *Cumming's* d_{unb} is de steekproefomvang reeds verdisconteerd in de effectgrootten van dit onderzoek (o.a. § 13.1). Wanneer gekozen zou zijn voor *weighted mean effect size* dan zouden de steekproefgrootten dubbel hebben meegewogen, vandaar de andere keuze.

Daarnaast ligt het voor de hand dat gecombineerde uitkomsten bij *weighted mean effect size* hoger zijn omdat positieve effectgrootten van de meest omvangrijke groepen het zwaarst wegen (Ellis, 2010, pp. 101-104). Bij grote meta-analytische studies is dergelijke assimilatie gerechtvaardigd, omdat de onderzoeker juist wil dat de belangrijkste studies - doorgaans die met de meeste deelnemers - het zwaarst wegen.

Desondanks doet dat niet ter zake bij dit onderzoek, omdat met name de methoden van de drie empirische studies, die deel uitmaakten van het *meervoudig onderzoek*, uiteenlopen. Maar, omdat zo'n gecombineerde waarde op eenvoudige wijze toch iets zegt over het effect - waar anders veel

beschrijvende woorden voor nodig zouden zijn - heb ik dus gekozen voor *simple mean effect size*. De gecombineerde effectgrootten (zie *Tabel 17*) zijn dus vooral bedoeld als indicatie, om daarmee aanvullend inzicht te bieden ten behoeve van de *praktische relevantie*.

13.5 Hoofdstukconclusie

Op basis van de resultaten van de *Praktische-relevantiestudie*, mag de conclusie zijn dat de opzet van dit meervoudig empirisch onderzoek geslaagd is om - via separate studies verkregen waarnemingen - meer inzicht te verkrijgen bij het vinden van goede antwoorden op de tweede onderzoeksvraag.

De praktische (educatieve) relevantie is over het algemeen relatief groot, wat inhoudt dat de gezamenlijke bevindingen conform de veronderstellingen zijn van de deelstudies (zie *Hoofdstuk 10, 11 en 12*). Anders gezegd: de hypothesen van die studies konden niet worden verworpen. Na alles te hebben overwogen, ben ik tot de volgende hoofdstukconclusie gekomen.

- De *Praktische-relevantiestudie* maakt het aannemelijk dat creativiteitstrainingen uit het curriculum inderdaad werkten voor de participerende *Bachelor of ICT*-studenten, met een navenant belang (*praktische relevantie*), maar dat desondanks het effect van de trainingen niet in alle gevallen even sterk was.

“These observations lead to a relatively unambiguous conclusion. Creativity training works”

Ginamarie Scott, Lyle E. Lertz & Michael D. Mumford (2004a, p. 370).

13.6 Antwoord op de tweede onderzoeksvraag

Heel *Deel II* van dit proefschrift stond in dienst van de eerste onderzoeksvraag: “Werkt creativiteitstraining die studenten *Informatie- en Communicatietechnologie* van Nederlandse hogescholen volgen vanuit het curriculum?” Op grond van de bevindingen van het totale meervoudige onderzoek (die over het algemeen in eenzelfde positieve richting wijzen), heb ik ook de tweede onderzoeksvraag met ‘Ja’ kunnen beantwoorden. Evenwel iets meer specifiek, is het antwoord op de tweede onderzoeksvraag dat creativiteitstraining over het algemeen, in positieve zin werkt bij *Bachelor of ICT*-studenten, ondanks dat het resultaat van de studies verschilde per training, en niet in alle gevallen even positief was. Feitelijk is daarmee ook het tweede deel afgerond.

Echter, omdat deze antwoorden niet voldoende verdieping tonen, heb ik, in lijn met *Deel II* (zie § 5.5), ook drie gedetailleerde conclusies geformuleerd. Daarnaast doe ik een concluderende aanbeveling aangaande de module *Creatieve Technieken* (CRT).

²⁹⁴ CRT vormde namelijk de controlegroep bij *Enquêtestudie*. Wel liet de *Enquêtestudie* eveneens de genoemde trend zien. Zo was bij de helft van de vragen de modus van de antwoorden op de zespuntsschaal gelijk aan of groter dan 4 (zie § 11.3). Daaruit mag de conclusie zijn dat een deel van de BICT-studenten, die CRT volgde, positief oordeelden. Als voorbeeld daarvan drie enquêtevragen die kenmerkend zijn voor het effect van CRT op hun creativiteit en de relevantie van creativiteitstraining voor BICT-studenten: “Door CRT ben ik creatiever gaan denken” (vraag 1); “Door CRT bedenk ik meer originele (oorspronkelijke) ideeën” (vraag 6); en “Creativiteitstraining is belangrijk voor HBO-informatica-studenten” (vraag 18). Zie § 11.2.2 en *Bijlage 6*, voor de beschrijving van alle enquêtevragen.

1. Het is aannemelijk dat creativiteitstraining uit het curriculum in positieve zin werkt voor *Bachelor of ICT*-studenten, met een navenant belang (*praktische relevantie*). En dat daarvan vooral sprake was wanneer zo'n training in handen was van een ervaren *facilitator*;
2. Ondanks die uitkomst was het effect niet in alle gevallen even sterk;
3. Zo vormde de CRT een negatieve uitzondering op dat positieve resultaat. Waarom CRT lager scoorde, viel niet precies af te leiden uit de empirische studies. Wel lijkt er steun voor te bestaan dat het effect, c.q. de werking, van een creativiteitstraining gevoelig is voor de ervaring van de *facilitator* van zo'n training, of juist voor het ontbreken daarvan.

Aanbeveling aangaande Creatieve Technieken (CRT)

- De relatief lage score van CRT mag tevens een indicatie zijn voor mijn advies deze lesmodule aan te passen, in ieder geval de verwachtingen ervan bij te stellen.

13.7 Overzicht van alle hoofdstukconclusies van Deel II

Mede voor het behoud van het overzicht, maar ook om bijvoorbeeld niet voortdurend te hoeven terugbladeren, en bovenal omdat de antwoorden op de eerste onderzoeksvraag uiteindelijk daarop zijn gebaseerd, volgt hier (in lijn met *Deel I*) weer een samenvatting (in 4 punten) van de daartoe doende hoofdstukconclusies van *Deel II*.

1. De resultaten van de *Interventiestudie (Hoofdstuk 10)* toonden (i) dat creativiteitstraining over het algemeen in positieve zin werkt bij *Bachelor of ICT*-studenten (vooral als een training in handen was van een ervaren, professionele *facilitator*). Ook al was (ii) het gemeten directe effect van creativiteitstraining gemiddeld minder sterk dan in het algemeen het geval is in het onderwijs (zie Hattie, *Hoofdstuk 3: Inleiding*). Dat mindere effect duidt (iii) op een mogelijke achterstand wat betreft creativiteit bij BICT-studenten, als mogelijke consequentie van het al eerder aangegeven *ontkoppelingsprobleem (§ 6.4)*. Die lagere score is van dusdanige betekenis dat, in de traditie van D.H. Cropley (zie § 6.2 tot § 6.4), (iv) gesproken mag worden van een *creativiteitsachterstand-probleem (creativity disadvantage problem)*. Alleen al op grond van die mindere creativiteit, maar uiteraard ook gezien de (zeer) hoge *praktische relevantie (§ 13.6)* - en dus het gegeven dat creativiteit van (groot) belang is voor BICT-studenten, binnen alle disciplines (zie § 5.5) - is het (v) raadzaam om creativiteitstraining op te nemen in ieder BICT-curriculum.
2. Bij de *Enquêtestudie (Hoofdstuk 11)* waren BICT-studenten over het algemeen van mening (ondanks dat die mening verschilde

per creativiteitstraining) dat: (i) creativiteitstraining in positieve zin werkt (weer met name als een training in handen was van een ervaren *facilitator*) – conform punt 1 en resultaten uit eerder onderzoek (zie o.a. *Hoofdstuk 3: Inleiding*); (ii) zij door creativiteitstraining zelf gemiddeld meer (creatieve) inspiratie hebben gekregen door de ideeën van anderen; en (iii) het voor iedereen relevant is om creativiteitstraining te volgen, inclusief zichzelf en alle andere HBO-studenten. Daarom is het raadzaam voor BICT-curriculumontwikkelaars – maar feitelijk voor alle leerlijnonwikkelaars in het onderwijs (van PO tot WO) – te overwegen om creativiteitstraining op te nemen in elk curriculum. Daarnaast is het opvallend dat CRT, ook bij de *Enquêtestudie* minder scoorde dan ASP, CD en LD. Het zou kunnen dat ook hier de onervarenheid van de studenten om te fungeren als *facilitator* van een creatieve sessie daarbij een rol heeft gespeeld. Tevens mag worden opgemerkt dat de *Enquêtestudie* een bijdrage levert aan onderzoek op het gebied van *Creative Self-Assessment* omdat daar (i) behoefte aan bestaat (§ 11.3.5); en (ii) omdat de enquête van deze studie kan worden hergebruikt (al dan niet in gedeelten) voor verder onderzoek.

3. De *Experts-studie (Hoofdstuk 12)* gaf indicaties dat creativiteitstraining werkte voor BICT-studenten bij het bedenken van creatievere ideeën, al waren er wederom uitzonderingen. Zo waardeerden creativiteitsexperts de ideeën van BICT-studenten die een creativiteitstraining volgden over het algemeen hoger op creativiteit dan de ideeën van dito studenten die dat niet deden. De bevindingen van de *Experts-studie* zijn in overeenstemming met punt 1 en 2, én wederom met de resultaten van eerder onderzoek (zie *Hoofdstuk 3: Inleiding*). Ook vormde CRT weer een uitzondering in negatieve zin, wat eveneens in lijn is met de punten 1 en 2.
4. De *Praktische-relevantiestudie (Hoofdstuk 13)*, oftewel de resultatensynthese dus van de drie voorgaande studies, maakt het aannemelijk dat creativiteitstraining uit het curriculum inderdaad positief werkt voor de participerende *Bachelor of ICT*-studenten, met een navenant belang (*praktische relevantie*), maar dat het effect per training niet overall even sterk was.

13.8 Tot slot van de tweede onderzoeksvraag

Nu, naast de antwoorden op de eerste onderzoeksvraag, ook de antwoorden op de tweede onderzoeksvraag bekend zijn, is het moment daar om de bevindingen over beide vragen samen te voegen. Dat doe ik in *Deel III (Hoofdstuk 14)*, het volgende en laatste onderdeel van dit proefschriftonderzoek.