Hermann Astleitner (Ed.)

# Intervention Research in Educational Practice

## Alternative Theoretical Frameworks and Application Problems

WAXMANN

Hermann Astleitner (Ed.)

# Intervention Research in Educational Practice

## Alternative Theoretical Frameworks and Application Problems

Waxmann 2020
Münster • New York

**Bibliographic information published by the Deutsche Nationalbibliothek**
The Deutsche Nationalbibliothek lists this publication in the
Deutsche Nationalbibliografie; detailed bibliographic data

Print-ISBN 978-3-8309-4197-2
Ebook-ISBN 978-3-8309-9197-7


Münster, Germany

www.waxmann.com
info@waxmann.com

Cover Design: Anne Breitenbach, Münster
Typesetting: MTS. Satz & Layout, Münster
Print: CPI Books GmbH, Leck

Printed on age-resistant paper,
acid-free according to ISO 9706


Printed in Germany

# Contents

## Contributors

Hermann Astleitner, A.Univ.-Prof. Mag. Dr.
https://www.uni-salzburg.at/erz/hermann.astleitner
Department of Educational Science
University of Salzburg
Erzabt-Klotz-Str. 1, 5020 Salzburg, Austria

Josef Eisner, Senior Lecturer Mag. Dr.
https://www.uni-salzburg.at/index.php?id=211343
Ulrike Greiner, Univ.-Prof. MMag. DDr.
https://www.uni-salzburg.at/index.php?id=57561
Franz Hofmann, Ao.Univ.-Prof. MMag. Dr.
https://www.uni-salzburg.at/index.php?id=211348
Michaela Katstaller, Mag. Dr.
https://www.uni-salzburg.at/index.php?id=60825
School of Education
University of Salzburg
Erzabt-Klotz-Str. 1, 5020 Salzburg, Austria

Nomy Dickman, Dr.
https://medicine.biu.ac.il/en/node/866
The Azrieli Faculty of Medicine
Bar-Ilan University
Ramat Gan 5290002, Israel

Gabriela Gniewosz, Dr.
https://diagnostik.sbg.ac.at/team/gniewosz/
Department of Psychology
University of Salzburg
Hellbrunner Straße 34, 5020 Salzburg, Austria

Stephanie Musow, MA
https://www.phsg.ch/de/team/ma-stephanie-musow
St. Gallen University of Teacher Education (PHSG)
Notkerstrasse 27
9000 St. Gallen, Switzerland

Alexander Naumann, Dr.
https://www.dipf.de/de/institut/personen/naumann-alexander
DIPF, Leibniz Institute for Research and Information in Education
Rostocker Straße 6, 60323 Frankfurt/Main, Germany

# Foreword

*The greatest difficulties lie where we are not looking for them.*
Johann Wolfgang von Goethe (1821/2000, p. 484)

*Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise.*
John Wilder Tukey (1962, p. 13–14)

This book aims to address special problems in "educational interventions" which are "measures that attempt to solve problems in the field of education" embedded within training, instructing, coaching, or counseling (Astleitner, 2010, p. 48). Such interventions are "purposeful actions" which "operate at the individual, family, organizational (e.g., school), neighborhood, regional, national, or other level. Interventions may be comprised of a single action or a cluster of actions" (Fraser, Richman, Galinsky, & Day, 2009, p. 5).

The selected problems in this book concern different areas of educational intervention research and are related to theoretical foundations as well as research designs and measurements. We focus on problems which are often overlooked in recent intervention research or excluded from scientific discourse although problem-related solutions could, above all, reduce the theory-practice gap or build links between basic research, applied research, and educational practice (e.g., Astleitner, 2003; Herber, 1998; Patry, 2001; Zumbach & Mandl, 2008). The given book should therefore be a supplement to existing standard approaches on educational intervention research (e.g., McBride, 2016; Phyne, Robinson, & Levin, 2005; Riley-Tillman & Burns, 2009; Stylianides & Childs, 2019).

Our problems and related studies were embedded in practical settings in the fields of higher education, teacher education, and teaching-learning-research.

A first major problem on theories in educational intervention research lies in distinguishing and handling successfully different types of theories from basic research, applied research, and educational practice. In particular, traditional theories suffer from their limited capacities in connecting different stages of human development with adaptive support strategies, in optimizing context sensitivity during implementation, and in handling the dark side of educational experiences. We will show, in chapter 1, how development-support, implementation-, and dark-side-theories could stimulate educational intervention research.

A second problem focuses on ineffective interventions in the field of teaching. Interventionistic researchers and practitioners need to know what kind of ineffective interventions and why ineffectiveness are given. Here, different types of teach-

ing malfunctions are identified. The contribution in chapter 2 also aims at formulating a macro- and a micro-theory about why teaching malfunctions occur.

Also, an important third problem in exploring effects patterns of educational interventions is given when a control group is missing because of practical, ethical, or other constraints. Here, we will demonstrate within an empirical study, in chapter 3, how a control group can be simulated and therefore have an impact on the conclusiveness of findings from quasi-experimental settings.

A fourth problem is about inconclusive or negative research findings. Even having a sophisticated theory- and research-based foundation within an educational intervention study will not guarantee that findings are significant and confirming given hypotheses. Sometimes, educational interventions improve some goal areas, but worsen others at the same time. In chapter 4, we will show how to provide theoretical and statistical analyses as well as a focus on alternative approaches in research design in order to handle such problems.

A fifth problem in educational intervention research is about the invalidity of measurements, especially when they are multidimensional in nature. Sometimes, it is necessary in educational intervention research to use measurement instruments without having the possibility to pretest them. Low validity of the measurement instruments might be the consequence. In chapter 5, we will show how statistical procedures on analyzing construct validity like confirmatory factor analysis can be used in order to test and handle these problems.

A sixth problem concerns a situation in which the correlations of pretests from different measurements are different for the posttests. When the correlations are different, then one might conclude that also the reliabilities and validities of measurements are different. When the reliability and validity of pretests are different from those on the posttests, then effect patterns in interventions (as differences between pre- and posttests) might be affected. In chapter 6, we will show and reflect a study in which such problems occurred.

A seventh problem is on valid interpretations and uses of test scores. In chapter 7, there is a focus on the question whether tests are appropriate or inappropriate to evaluate intended intervention goals. Here, the focus is on instructional sensitivity as a property of a test to capture effects of classroom interventions. The complex concept of instructional sensitivity appears to be very challenging for intervention research in educational practice and future research will show whether it can be beneficial.

Finally, an eight problem is on the question whether participants of an intervention are willing to give their best in a test on intervention effects. In order to answer this question, in chapter 8, a comprehensive overview of the state-of-the-art on test-taking motivation is given. Also, practical suggestions are depicted, how problems on test-taking motivation should be handled in educational intervention studies.

All problems are embedded in different types of research works, all of them can be found within the context of educational intervention research. There are concep-

tual analyses, theory building approaches, quasi-experimental intervention studies, or measurement validity studies. In case of the empirical studies, some of the mentioned problems were identified after the study was done. So, some of the problems were not first or starting points, but were found during the research process. Therefore, especially, within most of the empirical studies in this book, the orientation on the mentioned problems remains subdominant. This problem orientation is supplemented by innovative topics in the field of higher education resp. university teaching, teacher education and teaching-learning activities like alternative types of intervention theories, volitional competences, minimal guidance interventions, developmental measurement approaches, and so on.

Of course, these selected problems do not provide a complete list of problems in educational intervention research. We have focused on problems which are highly relevant for research activities in educational practice, but which are, at the same time, not prominently anchored especially in the field of teacher training or teaching-learning activities. It is also clear, that our way of handling such problems depends on our theoretical and methodological focus. This focus has multiple facets. These facets can be made explicit in the context of another important problem in educational intervention research concerning multiple testing.

## The Focus of This Book Exemplified by the Problem of Multiple Testing in Educational Intervention Research

In educational intervention research, it is often important to know effects of an intervention on multiple dependent variables. This situation implies that multiple statistical tests are used in one study. However, when conducting multiple tests, the problem of overall Type I error inflation (i.e. of rejecting the null hypothesis when it is true) occurs. Therefore, Alpha-levels should be adjusted on a lower level than the conventional .05 level. For such adjustments, many different methods can be found within the literature, some of them in different advanced versions and with available statistical software (e.g., Bretz, Hothorn, & Westfall, 2011; Shaffer, 1995). The most prominent procedure is an easy-to-use Bonferroni method in which each hypothesis is tested at an adjusted significance level of Alpha/n, whereas n is the number of hypotheses. Ignoring such methods has a strong impact in the scientific community: It could lead to false discoveries and result in wrong decisions on the effectiveness of interventions.

However, there are also critiques of such adjustment methods (Fink, McConnell, & Vollmer, 2014; O'Keefe, 2003): First, many correction methods dependent on the number of tests and therefore on the number of variables in a study. However, it is not always clear how many variables were in fact in the original study plan. Sometimes, researchers only report results of some pre-selected variables and not on all variables in a study. Second, there is some dispute on the question what counts as a distinct hypothesis. Some argue that all or many variables in a study are correlated,

and thus reflect only one single hypothesis. When there is only one hypothesis, then no adjustments of the Alpha-level are necessary. Third, reducing the Alpha-level leads to a problem concerning statistical power resp. Type II error. It reduces the chance of detecting a genuine (nonzero) effect. Forth, there is an inconsistent application of adjustment methods in a way that some researchers are applying such methods and some researchers in the same area of research are not using such procedures. As a consequence, the same results are sometimes statistically significant and sometimes confusingly non-significant.

Despite these problems, there is no doubt that Alpha-level adjustment is necessary especially in basic research, in which researchers have to be very conservative about accepting new evidence from experiments. However, in applied research as in the field of intervention research in educational practice, the situation is different in a way that higher Alpha-levels (e.g., $p > .05$) are accepted under certain circumstances. For example, Lipsey (1990, p. 39) wrote:

> A promising treatment might be investigated to determine if it has beneficial effects in some problem area. In such an applied research the implications of errors of inference may be quite different from those in basic research. To 'discover' that an applied treatment is effective when, in fact, it is not, does indeed mislead practitioners just as the analogous case misleads theoreticians. Practitioners, however, are often in situations where they must act as effectively as they can irrespective of the state of their formal knowledge, and it is unusual for them to use treatments and techniques of plausible but unproven efficacy. Moreover, demonstrably effective treatments for many practical problems are not easy to come by and candidates should not be to easily dismissed.

Educational intervention research is not only situated in the field of applied research, but also has often an exploratory goal focus. Within educational intervention research, this complex situation also leaded to the formulation of sophisticated guidelines for multiple testing. In such guidelines, it was suggested that "multiplicity adjustments are not required for exploratory analyses" (Schochet, 2008, p. 6), however that

> reports should explicitly state that exploratory analyses do not provide rigorous evidence of the intervention's overall effectiveness. Results from post hoc analyses should be reported as providing preliminary information on relationships in the data that could be subject to more rigorous future examination.

Others argued that in such exploratory analyses and related significance testing, *p*-values lose their meaning due to an unknown inflation of the Alpha-level and therefore research should focus on alternative criteria and methods instead of significance testing (e.g., simple graphic techniques) (e.g., Harlow, Mulaik, & Steiger, 1997). Recently, Rubin (2017, p. 272) stated that whether *p*-values lose their meaning depends on how we define the "family" of the error rate and argued that

> it is not necessary to lower the nominal Alpha-level when undertaking single tests of several different hypotheses. Exploratory analyses often involve many tests of this type. Consequently, Alpha-level adjustments are less necessary in exploratory analyses than would be the case if researchers adopted a multiple hypotheses approach to the familywise error rate. Again, it remains the case that the more hypotheses that a researcher tests, the greater probability that they will make a Type I error. However, this increased probability is distributed across the entire collection of hypotheses that are tested rather than localized to any one specific hypothesis. Consequently, it does not threaten the validity of any single test. If (a) researchers are interested in determining whether evidence supports or falsifies specific hypotheses rather than amorphous collections of hypotheses (i.e., universal null hypotheses), and (b) the probability that one hypothesis is true does not influence the probability that another hypothesis is true, then there is no need to adjust Alpha-levels for single tests of multiple hypotheses.

Finally, Trafimow et al. (2018) argued that manipulating the Alpha-level cannot cure significance testing and that the relative importance of Type I and Type II errors might differ across fields of research, studies, and researchers. Factors contributing to such differences include, for example, the clarity of theory, auxiliary assumptions, practical or applied concerns, or experimental rigor.

In this book, we reacted with the following strategies on these discussions in the field of multiple testing and other related problems:

a) We are clear about that our empirical studies represent applied research and have a significant exploratory orientation. In our studies, we tested certain interventions, or measurement approaches as well as related hypotheses for the first time. We will formulate and substantiate specific hypotheses, but most of them are in an early more or less exploratory state of theory development or scientific progress. Therefore, our findings are preliminary and have to be re-tested in future studies before they can be applied in educational practice.

b) As our empirical studies are situated in exploratory and applied settings, we will do no Alpha-level adjustments. We deliver exact *p*-values and therefore the possibility to evaluate their significance in case of simple Alpha-level adjustments. We assume that many of our findings will be non-significant if highly conservative Alpha-level adjustments would be made.

c) In educational practice in general, there are many factors that would decrease statistical power like low treatment integrity, small sample sizes, or floor or ceiling effects (Lipsey, 1990, p. 171). In the empirical studies in this book, there are some of these factors given, especially small sample sizes, due to uncontrollable situations in educational practice. In view of this problem, we do not apply Alpha-level adjustments because they would further decrease an already low power.

d) Of course, we are aware of the fact that multivariate testing would allow to give up Alpha-level adjustments resp. handle them effectively. However, we avoided multivariate testing due to methodological limitations of our studies. Multivariate testing, for example, with repeated measures multivariate analysis of variance would

Tab. 1: Differences between Basic Research, Applied Research, and Practice (based on: Astleitner, 2018, p. 149)

| Dimensions | Basic Research | Applied Research | Practice |
|---|---|---|---|
| Theory | Scientific theory | Technological theory, program theory | Subjective, implicit theory |
| Intervention | Experiment | Quasi-experiment, design-experiment | Case study |
| Significance | Statistical significance | Practical significance | Success |
| Validity | Validity of measurement and design | Usefulness | Problem solving |
| Measurement | Testing | Checking | Estimating |
| Explanation | Causality | Plausibility | Trial-and-error |

need adequate sample size, no univariate or multivariate outliers, multivariate normality, no multicollinearity, and others. On the one hand, such criteria are often not reached in small group studies in educational practice. On the other hand, results of multivariate testing are not significantly different from univariate testing in many cases (e.g., Keselman, Algina, & Kowalchuk, 2001). For example, within this book, in the study from Astleitner, Katstaller, and Greiner, five univariate analysis of variance produced the following 15 *p*-values (for five variables and related group, time, and group x time effects): .996, .815, .024, .303, .164, .073, .074, .079, .474, .336, .029, .796, .840, .042, and .421. A repeated multivariate analysis of variance produced the following *p*-values for the same tests: .594, .984, .039, .167, .156, .036, .085, .089, .446, .087, .016, .560, .755, .013, and .231. Only in 1 of 15 tests, the decisions on the hypotheses would be different: 0.073 (as non-significant) and 0.036 (as significant). Overall, *p*-values in both tests correlated with $r = .913$ ($p < .001$). The differences are mainly due to the fact that in multivariate analysis, all variables are included into the tests, what – due to missing cases – changes means and standard deviations as well as the resulting *F*-Tests.

e) Finally, we have to admit that our studies in this book suffer from typical methodological problems (e.g., small sample sizes) in the field of intervention research in educational practice. From a very strict experimental point of view, these shortcomings suggest to focus on descriptive results of our studies only. Therefore, in our studies, we do not only report and discuss statistical tests, but also descriptive statistics. An experimental psychologist in basic research might consider such descriptive information only, however, an educational researcher in applied and practical settings, would also consider results on hypotheses testing. Overall, we have to stress that there are significant differences between basic and applied research settings and that our studies are situated in applied resp. practical settings (e.g., Astleitner, 2013). Astleitner (2018, p. 149) outlined such differences (see Table 1). For example, in basic research, there is a focus on scientific theories (as classical if-then-statements). In applied research, we also focus on technological or program theories

(as if-do-statements). In practice, practitioners have subjective or implicit theories (as personal opinions).

## Case-based Learning

Having such differences in view of other areas of research and facing such problems in educational intervention research, lead to the conclusion that our findings should not be interpreted as effectiveness tests, but rather as possibilities for case-based learning in educational intervention settings. Zumbach, Haider, and Mandl (2008, p. 1) argued that such a perspective allows that learners "acquire knowledge through authentic problems from multiple perspectives, combining both foundations and application". To stimulate comprehensively and sustainably the acquisition of such a perspective should be an important goal of this book. This perspective lies in line with concepts of research like "transformative research" (e.g., Mertens, 2009) or "translational science" in which the major goal is to shorten and optimize the gap between knowledge production in science and application in practical contexts (e.g., Wehling, 2015). Therefore, the book is for researchers, instructors, designers, evaluators, or practitioners in many areas of education, especially on teacher education and teaching-learning activities. The book could be used within innovative scenarios on transdisciplinary educational approaches in which basic researchers, applied researchers, and practitioners work together to solve problems in educational practice (e.g., Ciesielski, Aldrich, Marsit, Hiatt, & Williams, 2017).

Multiple testing and related problems represent significant areas of concern in the field of educational intervention research. This book will show that there are many other problems to be found. However, despite all the difficulties and shortcomings in intervention research in educational practice, this book should be seen as an attempt to connect educational research and practice on a comprehensive and sophisticated basis.

Although we had a strong and farsighted scientific focus in our empirical studies, many of our results were non-significant or even negative. In our opinion, this is nothing unusual in highly complex and dynamic applied and exploratory settings, even if one can hardly find it in published work in professional journals. Results of this book stressed that educational activities, like, for example, teacher education might, at least partially, be ineffective and that this severe problem has to be taken seriously. In this book, we will present different explanations for such problems on a theoretical, design- as well as measurement-based orientation. Of course, all these explanations are preliminary and have to be tested in future research activities. Our book could represent a starting point for such attempts.

Of course, such an attempt needs the support of the scientific community. I would like to thank all the authors in this book for bringing their interest, creative ideas, and scientific skills to their contributions. For the different chapters in the book, the stated authors and co-authors were involved significantly in different stages of the research process (like planning, conducting, data analysis, and writing as

well as reviewing). Reviewing was conducted in a way that all co-authors delivered feedback to the texts and contributions of the other co-authors. The whole research was orientated on guidelines for good scientific practice and data protection from the University of Salzburg (retrieved from https://www.uni-salzburg.at). All participants in our studies were informed about the study goals, participated in the studies voluntarily and were able to cancel their participation at any time. Women were over-represented in most studies, but the generalizability of the results is not given anyway. All data was properly saved and statistical analyses were double checked. We also have evidence, that, as assumed before the interventions, control groups did not suffer from ethically questionable disadvantages in comparison to the intervention groups.

We have to thank all the participants in the studies. Thankfully, Jackie and Jörg Sams supported us in editing some texts in relation to the English language. We also thank for support by the Department of Educational Science and the School-of-Education of the University of Salzburg. We are also grateful for the help of student assistants in data entry and data management. We also want to thank Jörg Zumbach (University of Salzburg) for his critical, but true and constructive comments, especially on the problem of multiple testing. Thanks also to Hans-Jörg Herber and Jean-Luc Patry (University of Salzburg) who had and have a strong influence on my theoretical and methodological thinking. I was inspired by the shared work and encouraged to look at problems from multiple different perspectives what made me open and curious about future challenges in the field of educational intervention research. May that also be true for the readers of this book!

*Hermann Astleitner*
Salzburg, Spring 2020

## References

Astleitner, H. (2003). Praktische Signifikanz. *Journal für Lehrerinnen- und Lehrerbildung*, *3*, 48–53. Retrieved from https://www.studienverlag.at/zeitschriften/journal-fuer-lehrerinnen-und-lehrerbildung/

Astleitner, H. (2010). Methodische Rahmenbedingungen zur Entdeckung der Wirksamkeit von pädagogischen Interventionen [Methodological conditions for identifying the impact of educational interventions]. In T. Hascher & B. Schmitz (Eds.), *Pädagogische Interventionsforschung. Theoretische Grundlagen und empirisches Handlungswissen* (pp. 48–62). Weinheim: Juventa.

Astleitner, H. (2013, December). *Ist die Schulforschung naiv?* [Is school research naive?] Paper presented at the research colloquium of the School-of-Education, University Salzburg, Salzburg. Retrieved from https://www.uni-salzburg.at/fileadmin/multimedia/Erziehungswissenschaft/SOE2013_8.pdf

Astleitner, H. (2018). *Spezielle Verfahren sozialwissenschaftlicher Theorieentwicklung* [Special methods of theory building in social research]. Weinheim: Beltz Juventa.

Bretz, F., Hothorn, T., & Westfall, P. (2011). *Multiple comparisons using R*. Boca Raton: Chapman and Hall/CRC.

Ciesielski, T. H., Aldrich, M. C., Marsit, C. J., Hiatt, R. A., & Williams, S. M. (2017). Transdisciplinary approaches enhance the production of translational knowledge. *Translational Research*, *182*, 123–134. doi: https://doi.org/10.1016/j.trsl.2016.11.002

Fink, G., McConnell, M., & Vollmer, S. (2014). Testing for heterogeneous treatment effects in experimental data: False discovery risks and correction procedures. *Journal of Development Effectiveness*, *6*, 44–57. doi: https://doi.org/10.1080/19439342.2013.875054

Fraser, M. W., Richman, J. M., Galinsky, M. J., & Day, S. H. (2009). *Intervention research. Developing social programs*. New York: Oxford University Press.

Goethe, v. J. W. (1821/2000). *Werke* (Band 8). München: Deutscher Taschenbuch Verlag.

Harlow, L. L., Mulaik, S. A., & Steiger, J. H. (Eds.). (1997). *What if there were no significance tests?* Mahwah: Erlbaum.

Herber, H. J. (1998). Theorien und Modelle der Pädagogik, Psychologie und pädagogischen Psychologie – Annäherungsmöglichkeiten an ein komplexes Beziehungsproblem [Theories and models of pedagogy, psychology and pedagogical psychology – Approaching a complex relationship problem]. *Salzburger Beiträge zur Erziehungswissenschaft*, *2*, 41–101.

Keselman, H. J., Algina, J., & Kowalchuk, R. K. (2001). The analysis of repeated measures designs: A review. *British Journal of Mathematical and Statistical Psychology*, *54*, 1–20. doi: https://doi.org/10.1348/000711001159357

Lipsey, M. R. (1990). *Design sensitivity. Statistical power for experimental research*. Newbury Park, CA: Sage.

McBride, N. (2016). *Intervention research: A practical guide for developing evidence-based school prevention programmes*. Singapore: Springer. doi: https://doi.org/10.1007/978-981-10-1011-8

Mertens, D. M. (2009). *Transformative research and evaluation*. New York: Guilford.

O'Keefe, D. J. (2003). Colloquy: Should familywise alpha be adjusted? Against familywise alpha adjustment. *Human Communication Research*, *29*, 431–447. doi: https://doi.org/10.1111/j.1468-2958.2003.tb00846.x

Patry, J. L. (2001). Situation specificity of behavior: Triple relevance for research and practice in social research. *Salzburger Beiträge zur Erziehungswissenschaft*, *5*, 41–62. Retrieved from https://www.researchgate.net/profile/Jean_Luc_Patry

Phyne, G. D., Robinson, D. H., & Levin, J. (Eds.). (2005). *Empirical methods for evaluating educational interventions*. Burlington, MA: Elsevier Academic Press.

Riley-Tillman, T. C., & Burns, M. K. (2009). *Evaluating educational interventions. Single-case design for measuring responses to intervention*. New York: Guildford.

Rubin, M. (2017). Do p values lose their meaning in exploratory analyses? It depends how you define the familywise error rate. *Review of General Psychology*, *21*, 269–275. doi: https://doi.org/10.1037/gpr0000123

Schochet, P. Z. (2008). *Guidelines for multiple testing in impact evaluations of educational interventions* (Final report). Princeton: Mathematica Policy Research Inc. Retrieved from https://files.eric.ed.gov/fulltext/ED502199.pdf

Shaffer, J. P. (1995). Multiple hypothesis testing. *Annual Review of Psychology*, *46*, 561–584. doi: https://doi.org/10.1146/annurev.ps.46.020195.003021

Stylianides, G. J., & Childs, A. (Eds.). (2019). *Classroom-based interventions across subject areas. Research to understand what works in education*. Abingdon, New York: Routledge. doi: https://doi.org/10.4324/9781315170077

Trafimow, D., Amrhein, V., Areshenkoff, C. N., Barrera-Causil, C. J., Beh, E. J., Bilgiç, Y., et al. (2018). Manipulating the Alpha-level cannot cure significance testing. *Frontiers in Psychology*, *9*, 699. doi: https://doi.org/10.3389/fpsyg.2018.00699

Tukey, J. W. (1962). The future of data analysis. *The Annals of Mathematical Statistics*, *33*, 1–67. doi: https://doi.org/10.1214/aoms/1177704711

Wehling, M. (Ed.). (2015). *Principles of translational science in Medicine* (2nd ed.). Amsterdam: Elsevier.

Zumbach, J., & Mandl, H. (Eds.). (2008). *Pädagogische Psychologie in Theorie und Praxis: Ein fallbasiertes Lehrbuch* [Educational Psychology in theory and practice: A case-based textbook]. Göttingen: Hogrefe.

Zumbach, J., Haider, K., & Mandl, H. (2008). Fallbasiertes Lernen: Theoretischer Hintergrund und praktische Anwendung [Case-based learning: Theoretical background and practical application]. In J. Zumbach & H. Mandl, H. (Eds.), *Pädagogische Psychologie in Theorie und Praxis: Ein fallbasiertes Lehrbuch* (pp. 1–11). Göttingen: Hogrefe.

# PART 1.
# Theoretical Frameworks

# 1. Alternative Theoretical Frameworks for Educational Interventions[1*)]

*Hermann Astleitner*

Abstract: Educational intervention research is based on models of educational effectiveness. The goal of this contribution is to analyze the status of educational effectiveness models by focusing on theoretical concepts and criteria in basic and applied social sciences. This evaluation, which is based on a comprehensive exploratory review of literature, leads to the formulation of three alternative theoretical frameworks: Development-support-, implementation-, and dark-side-theories. Development-support-theories link developmental stages or competence levels (e.g., taxonomies of motivation) with adaptive strategies for establishing attainment-based support mechanisms (e.g., motivational tactics). Implementation-theories are about how interventions have to be designed for optimizing sensitivity to different contexts (like low- or high-quality educational scenarios). Dark-side-theories describe and explain the negative, non-transparent, or faulty facets of interventions (like trust and distrust as co-existing social realities for educational improvement). Discussions reflect on how to stimulate and develop intervention activities by using these alternative theoretical frameworks.

Educational interventions have the general goal to be effective in solving a given problem in educational contexts. Educational effectiveness is indispensable and depends on certain well-established criteria like statistical and practical significance, but also on theories or models of effectiveness.

## Challenges for Educational Effectiveness Theories

Existing theories on educational effectiveness which are essential in intervention research have to face multiple challenges arising from current developments in the field (e.g., Astleitner, 2010). Such challenges were identified in this contribution by conducting a comprehensive review of literature on educational effectiveness research and an evaluation of existing shortcomings. Such a review was not undertaken to find effect sizes or other indicators of effectiveness (as in traditional meta-analyses), but to get ideas for advancing theory building and research in the field.

1 *) This chapter is based on an unpublished paper presentation by Astleitner (2019).