

Contextual Factors in Treatment: A Meta-Analytic and an Experimental Approach

Inaugural Dissertation

Submitted to
the Department of Psychology at the University of Basel
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

By
Johanna Birkhäuer, M.Sc.

Basel, Switzerland
2017



Approved by the Faculty of psychology
at the request of

Prof. Dr. Rainer Greifeneder (Chair)

Prof. Dr. Jens Gaab (Referee)

Prof. Dr. Jana Nikitin (Co-Referee)

Basel, _____

Prof. Dr. Roselind Lieb (Dean)

Declaration of Authorship and Independence

The present articles were written in collaboration with the mentioned co-authors as partial fulfillment of the requirements for the degree of Doctor of Philosophy. No assistance from third parties was used and neither the author, nor the co-authors have published the articles elsewhere. Only the indicated tools are used and all quotes are marked

The following articles have been submitted for publication in context of this dissertation. Copies of the articles can be found in the appendix (A1-A3):

Birkhäuser, J., Gaab, J., Kossowsky, J., Hasler, S., Krummenacher, P., Werner, C. & Gerger, H. (2017). Trust in the health care professional and health outcome: A meta-analysis. *PLoS One*, 12(2), e0170988. doi:10.1371/journal.pone.0170988

Birkhäuser, J. & Gaab, J. (submitted). Experimental manipulation of interpersonal trust: an adaption of the trust game paradigm. Manuscript submitted for publication to *PLoS One*.

Tondorf, T., Kaufmann, L.-K., Degel, A., Locher, C., Birkhäuser, J., Gerger, H., Ehlert, U. & Gaab, J. (submitted). Placebo with psychotherapeutic meaning – a randomized controlled deceptive test of concept. Manuscript submitted for publication to *PLoS One*.

Basel, 2017

Johanna Birkhäuser

To my grandparents

Acknowledgments

Full of gratitude to my patient mentors Prof. Dr. Jens Gaab and Dr. Heike Gerger for their outstanding support and their enriching thoughts.

I would like to express my deepest thanks to my co-authors: Dr. Peter Krummenacher, Dr. Joe Kossowsky, Christoph Werner and Sebastian Hasler for their helpful comments and insights.

Special thanks go to Prof. Dr. Jana Nikitin for acting as the co-referee and Prof. Dr. Rainer Greifeneder for serving on the dissertation committee.

Finally, I wholeheartedly thank my family, friends and my colleagues and in particular Claudia Peter-Ruf, Michael Eichenberger, Bobby Xavier, Sibylle Meier und Cosima Locher for their great help and the funny and beautiful moments we shared in the last years.

Table of Contents

1. Abstract	3
2. Introduction.....	5
2.1 Context factors in medicine and psychotherapy.....	5
2.1 Testing for context factors in research.....	7
3. Aims of the thesis	9
4. Methods	10
4.1 Meta-analysis: Study 1. Trust and health outcome.....	10
4.1.2 Procedure and materials.....	10
4.1.3 Data analysis.....	10
4.2 Experimental studies: Study 2. Meaning – open/hidden evaluation	10
4.2.1 Participants.....	10
4.2.2 Procedure and materials.....	11
4.2.3 Data analysis.....	12
4.3 Experimental study: Trust induction.....	12
4.3.1 Participants.....	12
4.3.2 Procedure and materials.....	12
4.3.3 Data analysis.....	14
5. Summary of results.....	15
5.1 Trust in the health care professional and health outcome: A meta-analysis.....	15
5.2 Meaning – open/hidden evaluation	15
5.3 Trust induction.....	16
6. Discussion.....	17
6.1 Trust and health outcome	17
6.2 Meaning transformation by the use of an expressive writing paradigm.....	18
6.3 Trust induction by an adaptation of the trust game	19

6.4 Limitations	20
6.5 Conclusion and implications	22
7. References	26
7. Appendices	32

1. Abstract

Context factors have been hypothesized as being a core element in the medical treatment of the patient. Although their importance has extensively been described with regard to the medical and psychotherapeutic treatment, the empirical evidence of two factors, trust in the provider and meaning transformation in a treatment, remains vague. In order to study these two factors we used a two-folded approach: 1) A meta-analysis, quantifying the impact of trust on the health outcome and 2) the development of two experimental paradigms in order to provide a tool to systematically manipulate trust in provider and meaning transformation in order to test for their influence in research and clinical practice.

To quantify the impact of interpersonal trust in the clinical setting a meta-analysis was undertaken. Therefore, 5667 studies were screened. An overall mean effect as well as effect sizes for different subtypes of health outcomes were calculated. Patients reported higher treatment satisfaction, more beneficial health behaviors, less symptoms and a higher quality of life if they trusted their health care professional. With regard to objective health outcomes no direct link to trust was detectable with trust.

In a second step we tested the validity and feasibility of two placebo designs in order to study context factors in an empirical design. First, the effect of opposing therapy rationales was tested, using an open/hidden administration of an expressive writing intervention. The linguistic content analysis indicated plausibility of the experimental manipulation for the participants. Moreover, participants in the openly administered condition benefited from the expressive writing intervention in the long-run.

In the second experiment, we tested if the trust game, often used in economic psychology, might be a tool to induce trust for medical research. Therefore we used two fixed trust conditions. Randomly, half of the participants were allocated to the trust condition, implying a trustful participant-investigator-relationship, while the other half was distributed to

a non-trustful condition. We found that trust could systematically be manipulated, even controlling for emotions as anger or anxiety. The subjective rating of trust differed from the behavioral response of the participant.

Overall, these three papers try to contribute to a developing research on context factors. Mainly, I want to shed attention on two factors: trust in the provider and meaning transformation, which are less understood and tested in research. Rather than providing a theoretical debate on this topic, this dissertation aims to give an empirical insight and suggests two paradigms to elaborate research on this field. To quantify the mostly untested effects on trust and meaning transformation might be particular important in regard to a cost-benefit oriented health care policy.

2. Introduction

2.1 Context factors in medicine and psychotherapy

Patients are treated in a social and psychophysiological environment in which the interaction of the doctor and the patient is of high importance for the subjective as well as for physical health outcome (Spiegel & Harrington, 2008). In the health care setting, a situation that is mainly characterized by uncertainty, even well informed patients have to rely on the health care professional to work in their best interest (M. Calnan & Rowe, 2006). With regard to placebo research it has been shown that unspecific factors such as the patient-provider relationship as well as the treatment characteristics (e.g. the size of a tablet and especially the treatment rationale) seem to play a major role for the treatment outcome (Di Blasi, Harkness, Ernst, Georgiou, & Kleijnen, 2001). For psychotherapy more than for medical treatments, Frank (1986) has postulated that effects on the outcome have been found to result from the transformation of meaning that clients attach to their symptoms, relationships and life problems (Frank, 1986). Frank (1986) states that the key to healing is the elimination of the social isolation through supporting personal relationships and the transformed meaning patients attribute to their experiences and certain events. The transformation of meaning thereby involves an confiding relationship with a healer, a healing setting, a conceptual scheme or rationale, and procedures that both the healer and patient believe in and that involve active participation and positive expectations for change (Eysenck, 1994; Wampold, 2007a).

In this context, several authors have argued that the important component of meaning transformation is not to find the truth itself, but rather the formation of a new explanation that is plausible and adaptive - a coherent understanding of the world of the patient (Frank, 1986; Kazdin, 2007; Wampold, 2007b). This idea has been supported by a study showing that opposing breathing therapies result in the same effect as long as both therapies have a

comprehensive therapy rational (Sunyoung Kim, Eileen Wollburg, Walton T. Roth, 2012). One might conclude, that the rational therefore does not need to be of a specific content but rather has to be coherent in its own. Wampold supports this hypothesis by referring to the changing healing practices and their accompanying explanations, which are generated by the prevailing metaphysical zeitgeist of a particular society (Wampold, 2007b). However, empirical evidence for the effects of meaning transformation in psychotherapy is still lacking.

With regard to a second significant context factor, the effect of the patient provider relationship, Hippocrates described 400BC: “The patient, though conscious that his condition is perilous, may recover his health simply through his contentment with the goodness of the physician” (Di Blasi et al., 2001). Indeed, a study conducted by Kaptchuck and colleagues (2008) outlined that the effect sizes of combined outcome measures in a trial of placebo acupuncture in irritable bowel syndrome increased about 100% when the practitioner engaged in a warm and empathetic interaction with patients (Kaptchuk et al., 2008). Moreover, a meta-analysis by Kelley and colleagues (2014) found a small but significant effect of the relationship on the treatment outcome (Kelley, Kraft-Todd, Schapira, Kossowsky, & Riess, 2014). While the importance of the patient provider relationship is generally established in research, evidence and research is lacking for more specific factors, such as trust in the doctor.

The General Medical Council has declared the trustful relationship with the patient as a major duty of the doctor (General Medical Council). Besides its value per se, Blackstock and colleagues (2012) found that high trust in the clinical encounter was significantly associated with increased odds of antiretroviral adherence. They postulated the improvement of the trust level as a major implication for treatment (Blackstock, Addison, Brennan, & Alao, 2012). In accordance with these results, Hillen and colleagues (2011) concluded that a trustful relationship with the doctor resulted in a facilitated communication and a more efficient medical decision making process, a decrease of patient fear as well as an increased treatment

adherence and in the end resulting in a better health outcome of the patient (Hillen, de Haes, & Smets, 2011).

Although trust has been empirically studied in economics (Johnson & Mislin, 2011) and in neuroscience as well as in studies (Dunn & Schweitzer, 2005; Kosciak & Tranel, 2011) on the hormone oxytocin (Kessner, Sprenger, Wrobel, Wiech, & Bingel, 2013) no consensus on generally valid factors and mechanisms underlying the trust – health relationship have been found. To date, neither a global valid operationalization of the concept, nor an experimental approach of interpersonal trust can be found. For this dissertation the following definition of interpersonal trust has been fixed, although being aware that no “true” definition exists: “Trust in the physician is a patient's optimistic acceptance of a vulnerable situation, based on the expectation that the physician will act in the patient's best interest” (cited from Lee & Lin, 2009, p. 1060).

2.1 Testing for context factors in research

In order to systematically test and control for context factors in medical settings and in psychotherapy, specified paradigms need to be elaborated. As mentioned before experimental evidence drawn from prospective studies and randomized controlled trials are lacking for these factors.

With regard to trust, the trust game by Berg, Dickhaut and McCabe (1995) is a well-elaborated and replicated method in the domain of economic psychology, to measure trust in an experimental setting, including the sequential momentary exchange of two parties, with no predefined contract or agreement (Johnson & Mislin, 2011). There have been two variations of the trust game: 1.) the participant plays the trust game with a computer or 2.) with a randomly matched (sometimes anonymous) partner. However, to coherently test and study trust in a clinical setting, two changes might be necessary: 1.) trust needs to be induced systematically, by a standardized exchange and 2.) the trust induction must take place with a

doctor or therapist, which is not anonymous and not randomly matched, but has certain specific characteristics and acts in a particular professional frame. These changes, however, lead to a far fuzzier frame, potentially leading to different results from those in economic research. Therefore we tested whether it is possible to apply the trust game to induce trust in a participant by a predefined pattern in a non-anonymous investigator.

To test for meaning transformation in psychotherapy a similar issue arises. Whereas in biomedical research the "double blind technique" of randomized placebo controlled trials is considered to test for specific effects, this approach is not applicable to psychotherapy (Beecher, 1955a). Some attempts have been made to solve this problem (Chambless & Hollon, 1998; Grunbaum, 1981), all of them accompanying their own problems (Baskin, Tierney, Minami, & Wampold, 2003; C. L. S. H. J. Gaab, 2016; Kazdin & Wilcoxon, 1976). For instance with regard to systematic desensitization (SD), Locher et al. (2016) found that the restriction of therapeutic topics and of the emotional experience in the control condition appeared to be a prerequisite to find specific effects of the SD.

With regard to these results, the open versus hidden design might provide a tool to systematically test for the "meaning effects" of a treatment. In this paradigm an intervention is given openly (the participant knows when the drug is administered) and hidden (the participant does not know when the drug is given). This trial design has been used in non-pharmacological interventions (e.g. (Crum & Langer, 2007; Desharnais, Jobin, Cote, Levesque, & Godin, 1993) in order to specify effects resulting from the therapeutic meaning of a treatment (Moerman & Jonas, 2002). However, so far this paradigm it has not been tested for psychotherapeutic interventions. We therefore tested the feasibility of an open versus hidden design for psychotherapy.

3. Aims of the thesis

Context factors are key components in therapeutic as well as in medical interventions. However, their particular role remains unclear. Therefore, the current Ph.D. focuses on two context factors, interpersonal trust and meaning, and aims to measure their effects in the health care context as well as to develop a sophisticated and scientifically coherent design to study these factors in research. A meta-analysis was undertaken, to quantify the effect of interpersonal trust. In a second step, the two mentioned context factors were investigated with the help of placebo research, constructing two novel paradigms to change the meaning of a treatment and the trust level of the provider in a systematic manner. By this procedure, we aimed to provide a first experimental approach to empirically investigate these two factors.

4. Methods

The following section briefly introduces the study designs.

4.1 Meta-analysis: Study 1. Trust and health outcome

4.1.2 Procedure and materials

A systematic literature searches in CINAHL, Embase, MEDLINE and PsychINFO was conducted. We included studies that were (1) empirical and (2) reported quantitative data on the association between trust in the health care professional and health-related outcomes (exclusion of case vignettes, interviews, reviews, etc.), (3) took place in a health care setting, (4) were written in English or German, (5) were published journal articles and (6) measured interpersonal trust with a valid, reliable and established trust questionnaire. Two researcher independently extracted correlations between trust in the health care provider and health-related outcomes whenever reported.

4.1.3 Data analysis

Extracted data was converted to correlation with corresponding 95% confidence intervals (95% CI). As we expected considerable heterogeneity we performed a random-effects rather than a fixed-effects model. Data management and analysis (calculation of mean effect size) was performed by the help of Comprehensive Meta-analysis program (CMA).

4.2 Experimental studies: Study 2. Meaning – open/hidden evaluation

4.2.1 Participants

Overall, two samples were analyzed (completers N=112 and ITT N=126). Subjects were recruited amongst psychology students at two Swiss Universities (University of Zurich and University of Basel) through mailing lists and web postings at both universities. Inclusion criteria were (1) ages of 18 years and older, (2) absence of any mental disorder by self report, (3) not

receiving psychological or psychiatric or medical treatment in the last six month by self-report and (4) a Toronto Alexithymia Scale score below 54 (Kupfer, Brosig, & Brahler, 2000), since alexithymia has been shown to influence effects of expressive writing (Paez, Velasco, & Gonzalez, 1999; Solano, Donati, Pecci, Persichetti, & Colaci, 2003).

4.2.2 Procedure and materials

The study encompassed three intervention days (days 1 to 3) and four assessments (days 1, 4, 10, and 46, or baseline, post intervention, mid-term, and long-term follow-up, respectively). On intervention and assessment days, participants received an email with access to a webpage differing in content according to group assignment and study day. For intervention groups, on top of the outcome measure, the website contained an embedded video, showing a professional speaker explaining the rationale of the intervention. The instruction was equal for both intervention groups in terms of structure and format. The treatment rationale, however, differed in content between these groups. For the control group, the website contained the outcome only. The study was approved by the institutional review board of the Department of Psychology of the University of Basel.

Questionnaires

Influence on affect of the experimental conditions was repeatedly assessed with the Positive and Negative Affect Schedule (PANAS) (Krohne, Egloff, Kohlmann, & Tausch, 1996; Watson, Clark, & Tellegen, 1988). The PANAS contains two scales with overall 10 five-point items assessing positive (e.g. "interested", "proud") and negative affect (e.g. "upset", "ashamed").

In the intervention groups, linguistic content of written text, subjective rating of the severity of the reported traumatic experiences and the plausibility of their respective treatment rationale were employed to assess the validity of the intervention and experimental manipulation of treatment rationale. The linguistic content of the written text was analyzed

with *Linguistic Inquiry and Word Count* (Pennebaker et al., 2001; Wolf et al., 2008), which reliably quantifies words in a given text according to preset categories. Subjective rating of the severity of the traumatic experiences was assessed with a single sentence 5-point item (i.e. *In general, how distressing is the experience you have just written about for you?*) and plausibility of the treatment rationale was operationalized with a 5-point item for each group (causality group: *Writing about a traumatic experience influences my well-being*, reversed causality group: *My well-being influences how I write about a traumatic experience*).

4.2.3 Data analysis

A time by group by scale multivariate analysis of variance with subsequent time by group univariate analysis of variance for single PANAS scales was used to investigate differences between groups regarding treatment effects over time. Analysis of variance or chi-square tests were used to examine demographic and clinical variables at baseline. Statistical analysis was performed using IBM SPSS Statistics Version 22 and significance level was set to .05.

4.3 Experimental study: Trust induction

4.3.1 Participants

106 male participants were enrolled in the current study (mean age = 24.5 years, $SD = 4.7$ years). Participants were recruited by advertisement at the University of Basel, via the online recruitment system of the department of psychology (BAPS-Sona, <http://baps.sona-systems.com>) and on the website: markt.unibas.ch. We included (1) healthy (2) male students ((3)age 18- 40) with (4) sufficient German language skills.

4.3.2 Procedure and materials

Trust Game

The participants were told they were going to play a cognitive task where they could win some additional money, but instead played a trust game, adapted from Berg, Dickhaut and

McCabe (1995). Some changes were made in order to manipulate the participant's level of trust. Starting with an endowment of one start unit in each sequence, the participant could decide to send a particular portion of their endowment to the investigator. In total, five sequences were played. In sequence 1 and 2, the participant could send 25%, 50% or 75 % of their endowment; while in sequence 3, 4 and 5 also 0 or 100 % could be send. After the participant sent the money, the amount was tripled. The investigator (in the role of the trustee) in return passed the predefined amount of money back.

In the non-trust condition the investigator returned 50%, 0%, 25%, 0% and 0% of send amount respectively. With regard to the trust condition, based on the fairness heuristic, that trust should be facilitated when the decision yields maximum rewards for both parties, the experimenter always returned 50% in order to create a fair exchange (Burks, Carpenter, & Verhoogen, 2003). The amount passed by the sender was defined to capture trust.

Questionnaires

The State-Trait-Anxiety-Inventory (STAI) was employed before and after the trust game to measure anxiety. The STAI is a standardized self-report questionnaire, measuring anxiety both as state (X1) as well as a trait (X2). Both scales contain 20 items and a four-point ranking scale, ranging from 1 = almost never to 4 = almost always. Cronbach's alpha for both scales is found to be $\alpha = .90$ (Laux, 1981). Negative items were recoded.

Also, we used the State-Trait-Anger-Expression-Inventory (STAXI) before and after the trust game, which assesses five subscales of anger: state anger (10 items), trait anger (10 items), and anger expression (24 items). Items are rated on a four-point ranking scale from 1 = not at all to 4 = very much. A median Cronbach alpha coefficient of .87 for the STAXI has been reported, ranging from .75 to .82 for the different scales (Spielberger & Reheiser, 2009).

As a subjective measure of trust, participants were asked to indicate how trustful they perceived the investigator on a 10 cm adapted visual analogue scale (Lund, Vase, Petersen,

Jensen, & Finnerup). The intensity scale ranged from 0 (no trust) to 100 (high trust). Higher scores indicated higher trust.

4.3.3 Data analysis

We conducted an analysis of covariance (ANCOVA), including trust-condition as the fixed factor and the baseline trust measurement as the covariate. In a second step, the STAI- and STAXI-scores were included as covariates in order to control for the induction of anger or anxiety. To test whether the trust condition was also associated with differences in the amount sent by the participant, we calculated separate chi-square tests as an explanatory analysis.

Statistical analysis was performed using IBM SPSS Statistics Version 22. To verify the normality, linearity, (if AE) and homoscedasticity assumption of the residuals, scatter-, box- and QQ-plots were inspected and checked for outliers (leverage, cook's distance and mahalanobis distance). One subject was excluded. For all tests the significance level was set to .05.

5. Summary of results

5.1 Trust in the health care professional and health outcome: A meta-analysis

Results of Study 1 showed a significant association between trust in the health care professional and health outcome. However, results differed with regard to the outcome dimension, with small and non-significant correlations for objective and observer-rated outcomes and a moderate association with self-rated subjective outcomes. Patients reported to be more satisfied with treatment, to show more beneficial health behaviors, less symptoms and higher quality of life when they had more trust in their health care professional. With regard to the conducted meta-regressions, the association between trust and outcome was smaller in high quality studies. Interestingly, the observed association between trust and health appeared to be smaller in North America and Asia compared to Europe and Australia. Moreover, we found a larger association in studies that reported correlations and a lower association in studies that reported binary data.

5.2 Meaning – open/hidden evaluation

Both the intervention as well as the experimental manipulation proved valid and feasible, as indicated by the linguistic content analysis. The two groups did not differ substantially in their ratings of the treatment plausibility and indicated comparable severity of the traumatic experiences.

PANAS scores did not change and groups did not differ significantly over time for short and mid-term effects. However, groups differed in the long-term: After six weeks both intervention groups benefitted with higher positive affect in comparison to the no-treatment control group. With regard to negative affect, the causality group showed a decrease, while the reversed causality group did not benefit from the intervention.

5.3 Trust induction

Results of study 3 showed that the trust in the experimenter could be systematically manipulated by an adaptation of the trust game. Controlling for the baseline trust ratings, participants in the trust condition rated the experimenter as significantly more trustworthy at the post-rating than participants in the non-trust condition. Including anger and anxiety in the model this effect was attenuated but still detectable. For behavioral response, results showed a difference between trust conditions in round three, after the trustors' behavior was experimentally manipulated. Participant invested significantly more in the trust condition. For round four and five no difference could be detected between non-trust and trust. Moreover, analysis revealed a correlation between the subjective trust rating (post measure) and the investment in round three. No association was found for round four and five.

6. Discussion

The aim of this thesis was to study two context factors: trust in the provider and the meaning transformation in treatment. To provide an empirical estimate of the association between trust in the health care professional and health outcome we conducted a meta-analysis (study 1). With regard to psychotherapy research, the feasibility of an open/hidden paradigm (meaning transformation) (study 2), and of an experimental trust induction was tested (study 3), in order to provide a tool to systematically study those two context factors in health care research.

In study 1, we observed a significant association between trust in the health care professional and health outcome. Study 2 and study 3 presented two paradigms, which might provide a possible approach to further elaborate context factors in different settings. Our results show that the two context factors, trust and meaning transformation, have significant effects on the outcome of the treatment. Quantifying the effects of these factors might lead to a realization of their importance and as a consequence highlights the need for their preservation and enhancement in an evidence based and cost focused health care systems.

6.1 Trust and health outcome

We found a significant association between trust in the health care provider and health outcome, which quantifies the postulated importance of trust in health care (Benedetti, 2013; Brown & Calnan, 2012; Michael Calnan & Rowe, 2008; General Medical Council).

The non-significant association between trust and objective health outcomes in our meta-analyses might reflect an absence of such an association. However, since we observed a significant correlation between trust and self-rated subjective outcomes, which in turn have been associated with objective outcomes (Detmar, Muller, Schornagel, Wever, & Aaronson, 2002; Doyle, Lennox, & Bell, 2013; Simpson et al., 2006), it could also be hypothesized that a possible association between trust and objective outcomes depends on subjective variables,

such as adherence to medication or patient satisfaction with treatment (Giordano & Lindstrom, 2016; Street, Makoul, Arora, & Epstein, 2009).

The differential association for the subjective and objective outcomes —i.e. larger associations for the self-rated subjective outcomes and a small and non-significant association for objective outcomes—reflects previous findings on the two outcome dimension therefore was expected ((Beecher, 1955b; Hrobjartsson & Gotzsche, 2010; Schwarz & Buchel, 2015; Wechsler et al., 2011).

Subgroup analysis revealed a larger association in studies that reported correlations and a lower association in studies that reported binary data. With regard to the differences between the observed trust-outcome associations in Europe, North America, Australia, and Asia, our results are in part reflected in a recent survey on trust in physicians (Blendon, Benson, & Hero, 2014), Blendon and colleagues found that on a country list the US for instance near the bottom if asked whether they trust in their physician. However, with regard to satisfaction with the medical treatment the U.S. ranks near the top. In contrast, Switzerland is number one on both scales. This difference is in accordance with the correlations we found: a smaller correlation between trust in the provider and health outcome in North America than in Europe.

6.2 Meaning transformation by the use of an expressive writing paradigm

The results of the causality group reflected previous findings on expressive writing with a temporally delayed and medium-sized improvement in affect (Baikie, Geerligs, & Wilhelm, 2012; Beyer et al., 2014; Langens & Schuler, 2005, 2007a).

In line with the observed differences between the intervention groups, a recent within-subject placebo design observed a significantly higher analgesic effect for an overt in contrast to a covert administration of lidocaine injections (Lund et al., 2014) – showing that even medical parameters can be influenced by meaning. The observed effects of our trial support the

feasibility and utility of an open/hidden administration to disentangle effects of meaning transformation. In this regard, the beneficial effects on positive affect in both intervention groups could be seen as being caused by characteristic constituents of expressive writing, while the incidental treatment constituent of providing a treatment rationale is needed to obtain full effects, i.e. beneficial changes in both positive and negative affect.

6.3 Trust induction by an adaptation of the trust game

Despite the observed significant drop in both subjective as well as in behavioral trust measures after the trust manipulation, our results also indicate that a certain amount of trust resists even in the presence of unexpected behavior of the trustor. This resistance was apparent for both - the subjective experience of trust as well as for the behavioral level: Participants investigated more than 50% in each round independent from their condition. In line with our findings a study by Burnham and colleagues (2000) revealed that participants still sent money in the distrust condition. When looking at the history of trust research these results become clearer. Early social psychology has defined trust as a static phenomenon, where the participant either completely trusts or distrusts the counterpart (Rousseau, 1998). However, this view has been challenged when the development of trust (including phases of distrust as a part of trust development) was addressed for instance by Lewicki (Lewicki, Tomlinson, & Gillespie, 2006). Our results support the idea that trust is not an all-or-none phenomena but more a continuum, where both trust and distrust might appear together.

Also, we found that the subjective experience of trust did not correspond with the behavioral response in round four and five. Also, we did not find a difference between trust and non-trust condition with regard to investment. Moreover there was no correlation between trust rating and investment for round four and five. Glaeser and colleagues (2000) reported that behavior in the trust game was predicted by past trusting behavior and not by the responses of the participant on a – albeit attitudinal – trust survey. Based on this idea, it might hypothesized whether we judge a person as trustworthy might be independent on how we

react on this judgment, questioning at last partly whether a behavioral response can measure the perception of trustworthiness in appropriate way. Also, the lack of a correlation might represent a truly existing dissociation between the trust perception and the trust behavior. With regard to this hypothesis caution might be warranted when equalizing the investment (trust behavior) as a measure of trust experience in the trust game (Johnson & Mislin, 2011).

6.4 Limitations

Several limitations of our studies need to be considered. For the meta-analysis, there are indications that our overall results may be overestimated, since the significant Egger's test indicates a lack of small-scaled studies with non-significant correlations. This is supported by the fact that we found smaller associations in higher quality studies as well as in larger studies. The large fail-safe N in our analysis indicated, however, a very low risk for a non-significant overall association.

Second, the observed large associations between trust and subjective health outcome we found in our results could be explained by proximity between the two concepts (Hall, Dugan, Zheng, & Mishra, 2001; Thom, Ribisl, Stewart, Luke, & Phys, 1999). Data from a large-scale survey empirically confirms an association between trust in the health care professional and subjective health outcome (Croker et al., 2013). However, since different meta-analyses describe an enhanced risk of bias with regard to self-rated subjective outcomes, the found association in our analyses may result at least partially from an upward bias (Moustgaard, Bello, Miller, & Hrobjartsson, 2014; Wood et al., 2008). Third, we could not satisfactorily estimate the impact of possible moderators. Classification of patient characteristics was not possible, since most samples were mixed with respect to potentially relevant factors (e.g. ethnicity or disorder). However, statistical heterogeneity was small to moderate, which indicates only a small risk for the presence of strong moderators. Finally, it

needs to be noted that our analyses does not allow any predictions on causality of the observed association between trust in the health care professional and health outcome.

Also our experimental studies have distinct limitations, which need to be considered critically. In the open/hidden evaluation the participants in the reverse causality group may have known about the beneficial effects of writing. However, we opted against a direct inquiry, because asking participants about their assumptions regarding possibly positive effects of writing might prime such assumptions. Instead we assessed the plausibility of the respective condition rationale, which did not differ between the intervention groups. Second, in the open administration of the intervention we provided a description of expected positive effects of the intervention. Therefore, our instructions differed from those in Pennebaker (2000). However, also for the original instructions high positive expectancies were found (Langens & Schuler, 2007b). Finally, our hidden administration differed from similar approaches in medical research. In contrast to receiving a medication unaware of the administration, we provided a non-therapeutic rationale in order to ensure that both experimental conditions did not differ in their credibility (e.g. Langens & Schüler, 2007).

With regard to the trust game, we eliminated the possibility to capture changes across different rounds, because the subjective experience of trust was only measured before and after the trust game. By this procedure we tried to prevent again priming the participants to the topic of trust, thereby introducing a bias to our assessment (Légal, Chappé, Coiffard, & Villard-Forest, 2012). Second, we included, a dependent and an independent variable of trust in the same paradigm, which might have lead to circular nature of the experiment, since we induced and measured trust by the trust game. Since this problem is already embedded in the trust game, we aimed to counteract this by including the subjective trust rating as a second indicator variable. Finally, our trust induction might have been confounded by anger and anxiety, since our effect was attenuated when including anger and anxiety in the model. Understanding the particular nature of this association (anxiety, anger and in particular of

positive emotions and trust) and how these feelings interact might be particularly important with regard to patient centered care. Economic psychology has addressed this question already and might provide some insights also for health care research (e.g. Dunn & Schweitzer, 2005),

6.5 Conclusion and implications

Our meta-analysis highlights the need to differentiate between outcome dimensions in proposed trust models, since we found varying associations between trust and different type of outcomes. Also it needs to be investigated whether the positive impact of trust on subjective health, might in turn lead to improvements on objective outcomes (Wampold & Imel, 2015). We could not test for such an impact in our experiment as well as in the meta analysis – because of the supposed more complex nature of the phenomena including moderator variables like adherence to treatment, which was not addressable by our studies. Controlling and testing for the influence of demographic and cultural covariates might entangle the complex interaction. Also these complex interactions between trust and health outcome, including reverse causality, as well as the existence of moderators might be best addressable by prospective studies. For instance in a longitudinal study repeated trust measurements might provide a more precise insight on the trust development of the patient. Based on our result that a systematic trust induction is possible, not only the use of questionnaires but also methods to manipulate the trust level could give an answer on to causal questions on the trust - health outcome association.

A conceptual clarification of key variables would promote the debate on trust and allows the deduction of empirically testable hypotheses. Also the influence of distal factors (political, social and cultural variables) on the association between trust and outcome needs to be tested. For example Blendon and colleagues assume that the structures in which health care providers can take influence on health policy vary among countries, influences the trust of the

patient. As mentioned before Blendon and colleagues found a smaller correlation between trust in the provider and health outcome in North America than in Europe. The authors suggests, that Europe might have a more formal structures, in which physician leaders have the opportunity to negotiate with the government over issues, as payment and professional autonomy, which in turn might have an influence on the trust perception of the patient (Blendon et al., 2014). This direct assumed association between the health care system and trust of the patient, however, needs to be tested empirically.

It appears tempting to implement suggested predictors of trust as a cost-effective way to enhance trust in clinical care (Keating, Gandhi, Orav, Bates, & Ayanian, 2004; Ommen, Thuem, Pfaff, & Janssen, 2011). Training programs have been developed to enhance trust relations in clinical practice (e.g. communication skill training). So far these interventions have failed to show any effects on trust (McKinstry, Ashcroft, Car, Freeman, & Sheikh, 2006; Rolfe, Cash-Gibson, Car, Sheikh, & McKinstry, 2014). Bearing in mind that trust has been divided into confidence in professional competence and intentional trust, which implies that the patients feels that she/he is treated in her/his best interest (Michael Calnan & Rowe, 2008), it can be argued that with reference to the first part of trust (competence) skill trainings might improve health outcome, in terms of facilitating a competent and efficient treatment. Regarding intentional trust, these specific training programs might be of little benefit. On the contrary, training particular skills on how to act trustworthy might harm the authentic patient-provider relationship.

Moreover, in our paradigm the trust induction was found to be a complex phenomenon. For instance, the effect of the trust manipulation was influenced by anger and anxiety of the participants. Capra (2004) highlights that in the trust game identical logical structures (the same payoff) can be perceived differently depending on the participant's mood. Moreover, Eberl (2004) describe with regard to the trust game that participants check

for a sufficient emotional attachment to the experimenter. This is accordance with the idea of intentional or so called identification based trust, which is based on personal values and feelings of both parties (Roy J Lewicki & Wiethoff, 2006). In this context temporary emotions might form the basis for a longer trust development. Putting our results in the context of emotional attachment we suggest, that the trust game and the development of trust needs to be investigated in a more realistic social context, where the role of emotions and relationship towards a second person can be studied in a more systematic way.

Trust is essential when transforming meaning in treatment (Benedetti, 2013). In coherence with results on the context effects of trust, we found that meaning has an impact on treatment outcome. We were able to create a manipulation of the therapeutic meaning while keeping characteristics of the treatment constant. This open/hidden design might be promising for psychotherapy research and needs to be elaborated in order to test it in a clinical context. With regard to ethical considerations and based on our results common factor of meaning transformation should be transparent for the psychotherapist and the patient (J. Gaab, Bleuse, Locher, & Gerger, 2015). Frank (1986) stated that psychotherapy seeks to help patients to transform the meanings of their problems and symptoms by offering a narrative that explains the connection between symptoms and causes and thereby providing a way to overcome these obstacles. We found such an effect only for negative but not for positive affect. Based on these results, one might assume that meaning transformation could have a different impact on reducing negative symptoms than on enhancing positive health outcomes. To test this hypothesis I suggest studying the impact of meaning transformation with regard to different (negative and positive) health variables, as for instance clinical (e.g. number of symptoms, health behavior as adherence) or cognitive outcomes (e.g. reaction time, attention span).

To conclude, by the systematic manipulation and empirical testing of two context factors, namely trust in the provider and meaning transformation in treatment, it might not

only be possible to control for their effects in clinical studies, but also to systematically use them in order to improve the treatment of the patient.

7. References

- Baikie, K. A., Geerligs, L., & Wilhelm, K. (2012). Expressive writing and positive writing for participants with mood disorders: An online randomized controlled trial. *Journal of Affective Disorders, 136*(3), 310-319. doi:10.1016/j.jad.2011.11.032
- Baskin, T. W., Tierney, S. C., Minami, T., & Wampold, B. E. (2003). Establishing specificity in psychotherapy: A meta-analysis of structural equivalence of placebo controls. *Journal of Consulting and Clinical Psychology, 71*(6), 973-979. doi:10.1037/0022-006x.71.6.973
- Beecher, H. K. (1955a). The powerful placebo. *J Am Med Assoc, 159*(17), 1602-1606.
- Beecher, H. K. (1955b). The Powerful Placebo. *Jama-Journal of the American Medical Association, 159*(17), 1602-1606.
- Benedetti, F. (2013). Placebo and the New Physiology of the Doctor-Patient Relationship. *Physiological Reviews, 93*(3), 1207-1246. doi:10.1152/physrev.00043.2012
- Beyer, J. A., Lumley, M. A., Latsch, D. V., Oberleitner, L. M. S., Carty, J. N., & Radcliffe, A. M. (2014). Computer-based written emotional disclosure: the effects of advance or real-time guidance and moderation by Big 5 personality traits. *Anxiety Stress and Coping, 27*(5), 477-493. doi:10.1080/10615806.2013.868887
- Blackstock, O. J., Addison, D. N., Brennan, J. S., & Alao, O. A. (2012). Trust in primary care providers and antiretroviral adherence in an urban HIV clinic. *Journal of health care for the poor and underserved, 23*(1), 88-98.
- Blendon, R. J., Benson, J. M., & Hero, J. O. (2014). Public trust in physicians—US medicine in international perspective. *New England Journal of Medicine, 371*(17), 1570-1572.
- Brown, P., & Calnan, M. (2012). *Trusting on the edge: Managing uncertainty and vulnerability in the midst of serious mental health problems*: Policy Press.
- Burks, S. V., Carpenter, J. P., & Verhoogen, E. (2003). Playing both roles in the trust game. *Journal of Economic Behavior & Organization, 51*(2), 195-216. doi:Pii S0167-2681(02)00093-8Doi 10.1016/S0167-2681(02)00093-8
- Calnan, M., & Rowe, R. (2006). Researching trust relations in health care: conceptual and methodological challenges--introduction. *J Health Organ Manag, 20*(5), 349-358. doi:10.1108/14777260610701759
- Calnan, M., & Rowe, R. (2008). *Trust matters in health care*: McGraw-Hill Education (UK).

- Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology, 66*(1), 7-18. doi:Doi 10.1037//0022-006x.66.1.7
- Croker, J. E., Swancutt, D. R., Roberts, M. J., Abel, G. A., Roland, M., & Campbell, J. L. (2013). Factors affecting patients' trust and confidence in GPs: evidence from the English national GP patient survey. *BMJ Open, 3*(5). doi:10.1136/bmjopen-2013-002762
- Crum, A. J., & Langer, E. J. (2007). Mind-set matters: exercise and the placebo effect. *Psychol Sci, 18*(2), 165-171. doi:10.1111/j.1467-9280.2007.01867.x
- Desharnais, R., Jobin, J., Cote, C., Levesque, L., & Godin, G. (1993). Aerobic exercise and the placebo effect: a controlled study. *Psychosom Med, 55*(2), 149-154.
- Detmar, S. B., Muller, M. J., Schornagel, J. H., Wever, L. D. V., & Aaronson, N. K. (2002). Health-related quality-of-life assessments and patient-physician communication - A randomized controlled trial. *Jama-Journal of the American Medical Association, 288*(23), 3027-3034. doi:Doi 10.1001/Jama.288.23.3027
- Di Blasi, Z., Harkness, E., Ernst, E., Georgiou, A., & Kleijnen, J. (2001). Influence of context effects on health outcomes: a systematic review. *The Lancet, 357*(9258), 757-762.
- Doyle, C., Lennox, L., & Bell, D. (2013). A systematic review of evidence on the links between patient experience and clinical safety and effectiveness. *BMJ Open, 3*(1). doi:Art n E00157010.1136/Bmjopen-2012-001570
- Dunn, J. R., & Schweitzer, M. E. (2005). Feeling and believing: the influence of emotion on trust. *J Pers Soc Psychol, 88*(5), 736-748. doi:10.1037/0022-3514.88.5.736
- Eysenck, H. J. (1994). The outcome problem in psychotherapy: what have we learned? *Behav Res Ther, 32*(5), 477-495.
- Frank, J. D. (1986). Psychotherapy--the transformation of meanings: discussion paper. *J R Soc Med, 79*(6), 341-346. doi:10.1177/014107688607900611
- Gaab, C. L. S. H. J. (2016). When Do Placebos in Psychotherapeutic Research Work? A Systematic Review on the Example of Systematic Desensitization.
- Gaab, J., Blease, C., Locher, C., & Gerger, H. (2015). Go Open: A Plea for Transparency in Psychotherapy.
- General Medical Council. Duties of a doctor: The duties of a doctor registered with the General Medical Council. Retrieved from http://www.gmc-uk.org/guidance/good_medical_practice/duties_of_a_doctor.asp

- Giordano, G. N., & Lindstrom, M. (2016). Trust and health: testing the reverse causality hypothesis. *Journal of Epidemiology and Community Health, 70*(1), 10-16. doi:10.1136/jech-2015-205822
- Grunbaum, A. (1981). The placebo concept. *Behav Res Ther, 19*(2), 157-167.
- Hall, M. A., Dugan, E., Zheng, B. Y., & Mishra, A. K. (2001). Trust in physicians and medical institutions: What is it, can it be measured, and does it matter? *Milbank Quarterly, 79*(4), 613-+. doi:Doi 10.1111/1468-0009.00223
- Hillen, M. A., de Haes, H. C. J. M., & Smets, E. M. A. (2011). Cancer patients' trust in their physician-a review. *Psycho-Oncology, 20*(3), 227-241. doi:10.1002/pon.1745
- Hrobjartsson, A., & Gotzsche, P. C. (2010). Placebo interventions for all clinical conditions. *Cochrane Database of Systematic Reviews*(1). doi:ARTN CD00397410.1002/14651858.CD003974.pub3
- Johnson, N. D., & Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology, 32*(5), 865-889. doi:10.1016/j.joep.2011.05.007
- Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annu Rev Clin Psychol, 3*, 1-27. doi:10.1146/annurev.clinpsy.3.022806.091432
- Kazdin, A. E., & Wilcoxon, L. A. (1976). Systematic desensitization and nonspecific treatment effects: a methodological evaluation. *Psychol Bull, 83*(5), 729-758.
- Keating, N. L., Gandhi, T. K., Orav, E. J., Bates, D. W., & Ayanian, J. Z. (2004). Patient characteristics and experiences associated with trust in specialist physicians. *Archives of Internal Medicine, 164*(9), 1015-1020. doi:DOI 10.1001/archinte.164.9.1015
- Kelley, J. M., Kraft-Todd, G., Schapira, L., Kossowsky, J., & Riess, H. (2014). The Influence of the Patient-Clinician Relationship on Healthcare Outcomes: A Systematic Review and MetaAnalysis of Randomized Controlled Trials. *Plos One, 9*(4). doi:10.1371/journal.pone.0094207
- Kessner, S., Sprenger, C., Wrobel, N., Wiech, K., & Bingel, U. (2013). Effect of oxytocin on placebo analgesia: a randomized study. *Jama, 310*(16), 1733-1735.
- Koscik, T. R., & Tranel, D. (2011). The human amygdala is necessary for developing and expressing normal interpersonal trust. *Neuropsychologia, 49*(4), 602-611. doi:10.1016/j.neuropsychologia.2010.09.023
- Krohne, H. W., Egloff, B., Kohlmann, C. W., & Tausch, A. (1996). Untersuchungen mit einer deutschen Version der "Positive and Negative Affect Schedule (PANAS)". *Diagnostica, 42*, 139-156.

- Kupfer, J., Brosig, B., & Brahler, E. (2000). [Testing and validation of the 26-Item Toronto Alexithymia Scale in a representative population sample]. *Z Psychosom Med Psychother*, 46(4), 368-384.
- Langens, T. A., & Schuler, J. (2005). Written emotional expression and emotional well-being: the moderating role of fear of rejection. *Pers Soc Psychol Bull*, 31(6), 818-830. doi:10.1177/0146167204271556
- Langens, T. A., & Schuler, J. (2007a). Effects of written emotional expression: The role of positive expectancies. *Health Psychology*, 26(2), 174-182. doi:10.1037/0278-6133.26.2.174
- Langens, T. A., & Schuler, J. (2007b). Effects of written emotional expression: the role of positive expectancies. *Health psychology : official journal of the Division of Health Psychology, American Psychological Association*, 26(2), 174-182. doi:10.1037/0278-6133.26.2.174
- Laux, L. (1981). Das State-Trait-Angstinventar (STAI): Theoretische Grundlagen und Handanweisung.
- Lee, Y. Y., & Lin, J. L. (2009). The effects of trust in physician on self-efficacy, adherence and diabetes outcomes. *Social Science & Medicine*, 68(6), 1060-1068. doi:10.1016/j.socscimed.2008.12.033
- Légal, J.-B., Chappé, J., Coiffard, V., & Villard-Forest, A. (2012). Don't you know that you want to trust me? Subliminal goal priming and persuasion. *Journal of Experimental Social Psychology*, 48(1), 358-360.
- Lewicki, R. J., Tomlinson, E. C., & Gillespie, N. (2006). Models of interpersonal trust development: Theoretical approaches, empirical evidence, and future directions. *Journal of Management*, 32(6), 991-1022. doi:10.1177/0149206306294405
- Lewicki, R. J., & Wiethoff, C. (2006). Trust, trust development, and trust repair. *The handbook of conflict resolution: Theory and practice*, 92-119.
- Lund, K., Vase, L., Petersen, G. L., Jensen, T. S., & Finnerup, N. B. (2014). Randomised controlled trials may underestimate drug effects: balanced placebo trial design. *PLoS One*, 9(1), e84104. doi:10.1371/journal.pone.0084104
- McKinstry, B., Ashcroft, R. E., Car, J., Freeman, G. K., & Sheikh, A. (2006). Interventions for improving patients' trust in doctors and groups of doctors. *Cochrane Database of Systematic Reviews*(3). doi:ARTN CD00413410.1002/14651858.CD004134.pub2
- Moerman, D. E., & Jonas, W. B. (2002). Deconstructing the placebo effect and finding the meaning response. *Ann Intern Med*, 136(6), 471-476.

- Moustgaard, H., Bello, S., Miller, F. G., & Hrobjartsson, A. (2014). Subjective and objective outcomes in randomized clinical trials: definitions differed in methods publications and were often absent from trial reports. *Journal of Clinical Epidemiology*, *67*(12), 1327-1334. doi:10.1016/j.jclinepi.2014.06.020
- Ommen, O., Thuem, S., Pfaff, H., & Janssen, C. (2011). The relationship between social support, shared decision-making and patient's trust in doctors: a cross-sectional survey of 2,197 inpatients using the Cologne Patient Questionnaire. *International Journal of Public Health*, *56*(3), 319-327. doi:10.1007/s00038-010-0212-x
- Paez, D., Velasco, C., & Gonzalez, J. L. (1999). Expressive writing and the role of alexithymia as a dispositional deficit in self-disclosure and psychological health. *J Pers Soc Psychol*, *77*(3), 630-641.
- Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic Inquiry and Word Count (LIWC): LIWC2001*: Mahwah: Lawrence Erlbaum Associates.
- Rolfe, A., Cash-Gibson, L., Car, J., Sheikh, A., & McKinstry, B. (2014). Interventions for improving patients' trust in doctors and groups of doctors. *Cochrane Database of Systematic Reviews*(3). doi:ARTN CD00413410.1002/14651858.CD004134.pub3
- Rousseau. (1998). Introduction to the special topic forum (vol 23, pg 401, 1998). *Academy of Management Review*, *23*(4), 652-652.
- Schwarz, K. A., & Buchel, C. (2015). Cognition and the Placebo Effect - Dissociating Subjective Perception and Actual Performance. *Plos One*, *10*(7). doi:ARTN e013049210.1371/journal.pone.0130492
- Simpson, S. H., Eurich, D. T., Majumdar, S. R., Padwal, R. S., Tsuyuki, R. T., Varney, J., & Johnson, J. A. (2006). A meta-analysis of the association between adherence to drug therapy and mortality. *British Medical Journal*, *333*(7557), 15-18. doi:10.1136/bmj.38875.675486.55
- Solano, L., Donati, V., Pecci, F., Persichetti, S., & Colaci, A. (2003). Postoperative course after papilloma resection: effects of written disclosure of the experience in subjects with different alexithymia levels. *Psychosomatic medicine*, *65*(3), 477-484.
- Spiegel, D., & Harrington, A. (2008). What is the placebo worth? *British Medical Journal*, *7651*, 967.
- Spielberger, C. D., & Reheiser, E. C. (2009). Assessment of Emotions: Anxiety, Anger, Depression, and Curiosity. *Applied Psychology-Health and Well Being*, *1*(3), 271-302. doi:10.1111/j.1758-0854.2009.01017.x

- Street, R. L., Makoul, G., Arora, N. K., & Epstein, R. M. (2009). How does communication heal? Pathways linking clinician–patient communication to health outcomes. *Patient Education and Counseling*, 74(3), 295-301.
- Thom, D. H., Ribisl, K. M., Stewart, A. L., Luke, D. A., & Phys, S. T. S. (1999). Further validation and reliability testing of the Trust in Physician Scale. *Medical Care*, 37(5), 510-517. doi:Doi 10.1097/00005650-199905000-00010
- Wampold, B. E. (2007a). Psychotherapy: "The" humanistic (and effective) treatment. *American Psychologist*, 62(8), 857-873. doi:Doi 10.1037/0003-066x.62.8.857
- Wampold, B. E. (2007b). Psychotherapy: the humanistic (and effective) treatment. *Am Psychol*, 62(8), 855-873. doi:10.1037/0003-066X.62.8.857
- Wampold, B. E., & Imel, Z. E. (2015). Great Psychotherapy Debate: The Evidence for What Makes Psychotherapy Work, 2nd Edition. *Great Psychotherapy Debate: The Evidence for What Makes Psychotherapy Work, 2nd Edition*, 1-323.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *J Pers Soc Psychol*, 54(6), 1063-1070.
- Wechsler, M. E., Kelley, J. M., Boyd, I. O. E., Dutille, S., Marigowda, G., Kirsch, I., Kaptchuk, T. J. (2011). Active Albuterol or Placebo, Sham Acupuncture, or No Intervention in Asthma. *New England Journal of Medicine*, 365(2), 119-126.
- Wolf, M., Horn, A., Mehl, M., Haug, S., Pennebaker, J. W., & Kordy, H. (2008). Computergestützte quantitative Textanalyse: Äquivalenz und Robustheit der deutschen Version des Linguistic Inquiry and Word Count [Computer-aided quantitative text analysis: Equivalence and reliability of the German adaptation of the Linguistic Inquiry and Word Count]. *Diagnostica*, 1, 85-98.
- Wood, L., Egger, M., Gluud, L. L., Schulz, K. F., Juni, P., Altman, D. G., Sterne, J. A. C. (2008). Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *British Medical Journal*, 336(7644), 601-605. doi:10.1136/bmj.39465.451748.AD

7. Appendices

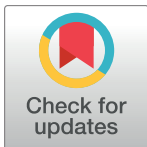
RESEARCH ARTICLE

Trust in the health care professional and health outcome: A meta-analysis

Johanna Birkhäuser^{1*}, Jens Gaab¹, Joe Kossowsky^{1,2,3}, Sebastian Hasler¹, Peter Krummenacher⁴, Christoph Werner¹, Heike Gerger¹

1 Clinical Psychology and Psychotherapy, Department of Psychology, University of Basel, Basel, Switzerland, **2** Department of Anesthesiology, Perioperative and Pain Medicine, Boston Children's Hospital/Harvard Medical School, Boston, Massachusetts, United States of America, **3** Program in Placebo Studies and the Therapeutic Encounter, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts, United States of America, **4** Collegium Helveticum, University of Zurich and ETH Zurich, Zurich, Switzerland

* j.birkhaeuer@unibas.ch



OPEN ACCESS

Citation: Birkhäuser J, Gaab J, Kossowsky J, Hasler S, Krummenacher P, Werner C, et al. (2017) Trust in the health care professional and health outcome: A meta-analysis. PLoS ONE 12(2): e0170988. doi:10.1371/journal.pone.0170988

Editor: Urs M. Nater, University of Marburg, GERMANY

Received: September 21, 2016

Accepted: January 13, 2017

Published: February 7, 2017

Copyright: © 2017 Birkhäuser et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: Johanna Birkhäuser's contributions to this study were supported by the Swiss National Science Foundation (<http://www.snf.ch>), grant project (POBSP1_155446). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: I have read the journal's policy and the authors of this manuscript have the

Abstract

Objective

To examine whether patients' trust in the health care professional is associated with health outcomes.

Study selection

We searched 4 major electronic databases for studies that reported quantitative data on the association between trust in the health care professional and health outcome. We screened the full-texts of 400 publications and included 47 studies in our meta-analysis.

Data extraction and data synthesis

We conducted random effects meta-analyses and meta-regressions and calculated correlation coefficients with corresponding 95% confidence intervals. Two interdependent researchers assessed the quality of the included studies using the Strengthening the Reporting of Observational Studies in Epidemiology guidelines.

Results

Overall, we found a small to moderate correlation between trust and health outcomes ($r = 0.24$, 95% CI: 0.19–0.29). Subgroup analyses revealed a moderate correlation between trust and self-rated subjective health outcomes ($r = 0.30$, 0.24–0.35). Correlations between trust and objective ($r = -0.02$, -0.08–0.03) as well as observer-rated outcomes ($r = 0.10$, -0.16–0.36) were non-significant. Exploratory analyses showed a large correlation between trust and patient satisfaction and somewhat smaller correlations with health behaviours, quality of life and symptom severity. Heterogeneity was small to moderate across the analyses.

Conclusions

From a clinical perspective, patients reported more beneficial health behaviours, less symptoms and higher quality of life and to be more satisfied with treatment when they had higher

following competing interests: the authors declare that they have published on the topic of context effects before. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

trust in their health care professional. There was evidence for upward bias in the summarized results. Prospective studies are required to deepen our understanding of the complex interplay between trust and health outcomes.

Introduction

Patients' trust in their health care professional is central to clinical practice [1, 2]. The General Medical Council states that "(p)atients must be able to trust doctors with their lives and health" and that maintaining trust is one core guidance for physicians [3]. Similar obligations are part of codes of conduct for other health care professionals such as nurses [4] or psychotherapists [5]. Patients have to trust their health care professionals to work in their best interest and outcome [6]. In this regard, trust in the health care professional has been suggested to be the foundation for effective treatments [7, 8] and fundamental for patient-centered care [9].

Besides such a deontological obligation for trust theoretical models describe mechanisms on how trust may influence health outcomes [7, 10–12]. Some of those conceptualize trust in the health care provider in relation to the patient-clinician relationship, which has previously been shown to be significantly associated with health outcomes across 13 RCTs [13]. Therefore, the question arises whether trust in the health care professional is as well associated with patients' health.

Empirical evidence regarding this question comes from a growing number of studies that report correlations between trust measures and patients' health outcome. In the different studies the health outcomes encompass different dimensions, such as objectively measured indicators (e.g. CD4 cell counts) [14], clinical observations (e.g. clinical diagnoses) [15], and patients' subjective self-ratings (e.g. patient satisfaction) [16]. The association between trust and health outcome has been found to differ across individual studies. For instance, in a sample of patients with diabetes, trust in the health care professional was found to be positively related to objective and subjective health outcomes (glycemic control, health-related quality of life, and patient satisfaction) [12]. In contrast, there was no significant association between trust in the health care professional and subjective outcomes (blood pressure control) in patients with hypertension [17]. In the absence of a systematic and comprehensive summary of the available evidence the variation in the observed health outcomes and in disorders complicates conclusions regarding the association between trust and health.

The empirical confirmation of the assumed association between trust and health outcome would strengthen the—so far—ethically derived claims for trustful and patient-centered relationships in clinical settings [2, 18, 19]. Therefore, a comprehensive and differentiated summary of the available evidence is needed. We conducted a systematic review with meta-analysis of observational studies in order to (1) estimate the overall association between trust and health outcome and to (2) investigate whether the strength of such an association depends on the type of health outcome. We controlled for a potential impact of study methodology and design on the association between trust and health outcome.

Methods

Data sources and searches

We conducted a systematic literature search in bibliographic databases (CINAHL, Embase, MEDLINE and PsycINFO; see [S1 File](#), which summarizes the applied search strategies). All records were transferred to EndNote (EndNote X7 Thomson Reuters, USA), where duplicates

were eliminated and all titles and abstracts were screened for inclusion and exclusion criteria. Records were excluded if they clearly did not meet our inclusion criteria (J.B.). Two researchers then independently reviewed the full text of the records that were considered potentially relevant during title and abstract screening (J.B. and C.W.). Ambiguities were resolved by consensus or by consulting a third researcher (H.G.). If only the title, the author or journal names were provided by the electronic database search, we contacted the authors or searched the journal archives manually in order to check for eligibility. Records were excluded if we were unable to obtain the full-text (Fig 1).

Study selection

We included studies that (1) reported quantitative data on the association between trust in the health care professional and health outcomes (exclusion of case vignettes, interviews, reviews, etc.), (2) took place in a health care setting, were written in English or German, (4) were published journal articles and (5) measured interpersonal trust (e.g. trust in the nurse, physician, GP, psychiatrist) with a valid, reliable and established trust questionnaire (i.e. included a reference to a published article which used the respective trust questionnaire; self-created, adapted and single-item questionnaires were excluded).

To qualify as health outcome the dependent variable needed to relate to at least one of the following health dimensions: (a) aetiology of the health problem (e.g. physiological measures, such as CD4 cell count, HbA1c), (b) symptoms (e.g. depression, worries, mood), (c) treatment-related

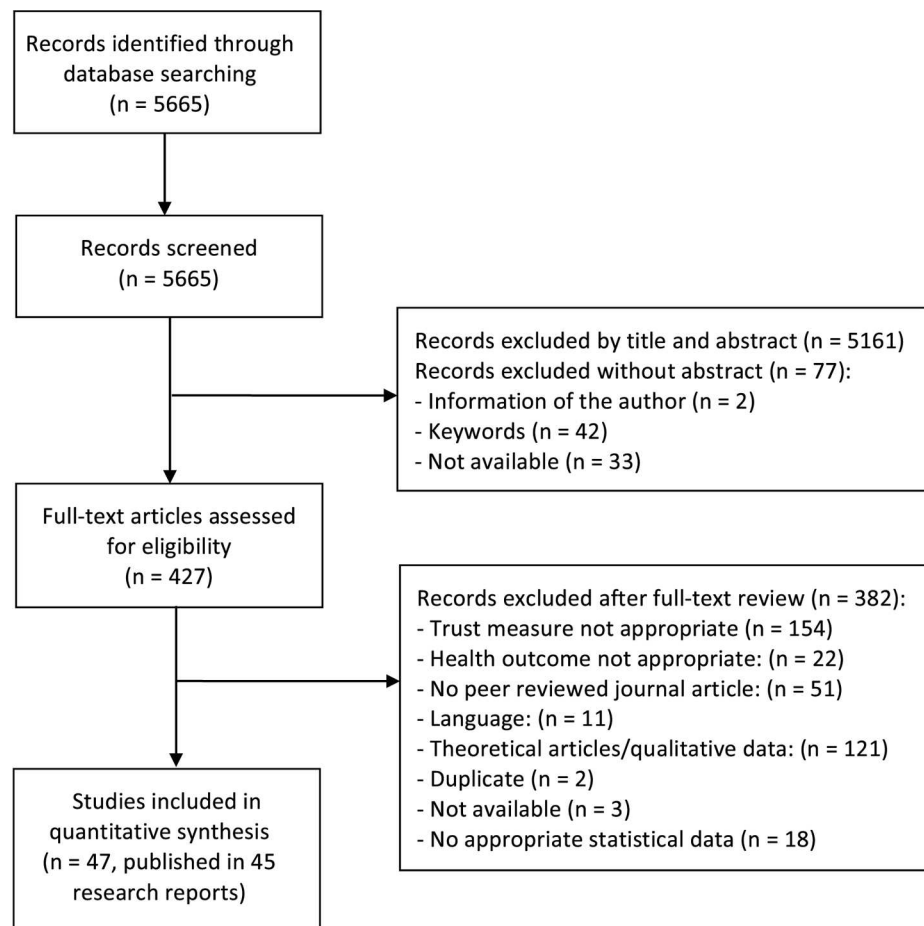


Fig 1. Study selection procedure.

doi:10.1371/journal.pone.0170988.g001

indicators (e.g. adherence, health promoting lifestyle, satisfaction with treatment), or (d) consequences of being in treatment (e.g. patient satisfaction, health-related quality of life, functional level, overall health, cognitive and emotional change). In cases of uncertainty (e.g. online search behaviour) we required that the authors provide at least one reference to an article that showed an association between a particular outcome and one of the previously outlined health dimensions. Moreover, the outcome needed to have a clear direction in order to distinguish between positive versus negative health outcome (e.g. increase in CD4 cell count, and decrease on a scale measuring depression symptom severity indicate improvement). We excluded outcomes whenever it remained unclear whether an increase or decrease on the respective measure was a beneficial outcome for the patient (see [S1 Table](#) for a list of the included health outcomes).

Data extraction and quality assessment

Two researchers (J.B. and S.H.) independently extracted correlations between trust in the health care provider and health outcomes whenever reported. If correlations were not reported, we used the available binary data (i.e. numbers / proportions of events) or the provided odds ratios in order to estimate the association between trust in the health care provider and health outcome. Binary data and odds ratios were transformed into log odds ratios, then into Fisher's Z and finally into correlations. If data from two studies were reported in one publication, we extracted data from each study independently.

We extracted the diagnoses in the patient sample, the country in which the study was conducted, the duration of treatment and characteristics of the trust scale. For quality assessment of the included studies we used the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist with a maximum of 22 points [20].

Health outcomes were clustered into objective (e.g. blood pressure, HbA1C, BMI), observer-rated (e.g. diagnosis by a professional) as well as self-rated subjective outcomes. The latter were divided into self-reported health behaviours (e.g. medication adherence, screening behaviour, health promoting lifestyle, online search behaviour) and health-related subjective experiences (e.g. patient satisfaction, health-related quality of life, pain-related anger, depression, worries).

Following a thorough training, data extraction and coding was conducted independently by two researchers (J.B. and S.H.) using a standardized form (Microsoft Office Excel 2007). Discrepancies between the two coders were resolved in face-to-face discussions or by consulting a third researcher (H.G.) when no consensus could be found.

Data synthesis and analysis

We calculated correlation coefficients in order to estimate the association between trust in the health care professional and health outcome with the corresponding 95% confidence interval (CI). With regard to the magnitude of effect sizes, we interpreted correlation coefficients in the order of 0.10 as "small", those of 0.30 as "moderate," and those of 0.50 or higher as "large" [21]. Whenever data was reported for subgroups of study participants in one publication, we calculated the weighted mean correlation [22]. If the N was missing in the table of analysis, we used the N of the descriptive statistics, ignoring the possibility that not all study participants may have contributed data for the correlational analyses. If only subscales of a composite health outcome measure were reported, we coded the available subscale data according to our classification of health outcomes. However, we preferred total scores over subscale data, if available. When more than one health outcome was reported in a publication, we combined the data from different outcome measures so that each study contributed only one estimate per analysis to avoid dependencies in the data.

In a primary analysis, we calculated an overall estimate across all health outcomes. Then, we calculated separate estimates for objective, observer-rated and self-rated subjective outcomes. We further divided subjective self-rated outcomes into health behaviours and subjective experiences and estimated the impact of potential moderators (i.e. study quality, duration of the treatment, the country in which the study was conducted, and the applied trust questionnaire). In an exploratory sub-analysis we divided subjective experiences into patient satisfaction, quality of life and symptom-related outcomes and conducted individual subgroup-analyses. For the subgroup and moderator analyses including categorical predictors we conducted stratified meta-analyses; in case of continuous outcomes we conducted meta-regressions.

We explored the presence of small study bias and publication bias by assessing funnel plot asymmetry (i.e. whether negative or non-significant findings are missing) with a regression test [23]. We inspected the Egger's regression coefficient rather than the Begg's correlation, since the power for this test is higher [24]. We calculated a fail-safe N, which determines the number of unretrieved studies with a null-finding that would bring the pooled estimate to zero [25]. Finally, we applied the trim and fill method, which adjusts the association between trust in the health care professional and health outcome for missing studies using the random effects model and adjusting for studies missing at the left side of the mean [26]. We used Comprehensive Meta-Analysis, version 2.0 (available at www.meta-analysis.com) for all meta-analyses and subgroup analyses, and STAT 13.1 for the meta-regressions. We used a two-sided *P*-value to test for statistical significance.

We applied random effects rather than a fixed effect model, since the included studies were expected to be heterogenous in several respects. To evaluate heterogeneity between studies, we examined τ^2 , which is an estimate of the variance among true effect sizes. τ (square root of τ^2) represents the standard deviation of the distribution underlying the included trials assumed to be a random sample. Higher τ^2 -values indicate greater variability between studies than would be expected by chance. Based on the definition of small, moderate, and large effect size estimates according to Cohen in 1988 [21] we interpreted τ^2 as follows: $\tau^2 = (0.2/2)^2 = 0.01$ was considered to represent low heterogeneity, $\tau^2 = 0.06 [(0.5/2)^2]$ moderate heterogeneity, and $\tau^2 = 0.16 [(0.8/2)^2]$ high heterogeneity between studies. As a measure of heterogeneity, τ^2 has been shown to be independent of the number of studies and patients included in a meta-analysis (i.e. no increase with large numbers of studies or large sample sizes) [27].

Results

Descriptives of included studies

We included 47 studies that were published in 45 reports (see [S1 Table](#) for descriptive information and the references of the included studies) with 34 817 participants (median: 200 participants, range: 24 to 8392). The procedure of study selection, including reasons for exclusion after full-text review, is shown in [Fig 1](#). Studies were conducted in Asia (2), Europe (6), North America (34), and Australia (2). Two studies did not specify the country, and one study was conducted in more than one country. In 24 studies trust was measured by the Trust in Physician Scale [28], six studies used the Wake Forest Trust Scale [29], four studies the Trust Scale of the Primary Care Assessment Survey [30], three studies applied the Trust Scale of the Illness Concept Scale [31], two used the Consumer Assessment of Healthcare Providers and Systems (CAHPS) Cultural Competence [32], and two used the Trust Scale of the Cologne Patient Questionnaire [33]. Six trust questionnaires were only used in one study (see [S1 Table](#) for details). 33 studies reported correlations between trust in the health care professional and health outcome and 15 studies reported binary data or odds ratios. Most studies did not report

the order of assessment of trust and outcome. Four studies specified that the trust measurement preceded the outcome assessment. No study reported on a reverse order of data collection (i.e. outcome assessment preceded the trust assessment). If no information was given, we assumed that trust and outcome were assessed at the same time-point. 15 studies reported objective health outcomes; two studies reported observer-rated outcomes and 42 studies reported subjective self-rated outcomes. Most studies included participants with chronic and multiple health complaints (see [S1 Table](#)). Six studies defined the duration of treatment with a median of 1.7 months (range: 10 days to 5 years). Two studies reported a mean of 2.6 and 2.7 visits in the study sample. Study quality ranged from 7.5 to 18.5 STROBE points (median: 15 mean: 13.5). Studies were published between 1981 and 2013 (median: 2009).

Association between trust and health outcome: Primary, secondary and exploratory analyses

Across all outcomes, we found a small to moderate correlation between trust and health outcome ($r = 0.24$, 95% CI: 0.19 to 0.29) based on all 47 studies ([Table 1](#) and [S2 File](#), which shows all forest plots). Heterogeneity between studies was low to moderate.

Stratified analyses revealed small and non-significant correlations in case of objective outcomes and observer-rated outcomes ($r = -0.02$, -0.08 to 0.03 and $r = 0.10$, -0.16 to 0.36 respectively; see [S2 File](#) for the respective forest plots), as well as a moderate correlation with regard to subjective self-rated outcomes ($r = 0.30$, 0.24 to 0.35). Heterogeneity was small to moderate.

In a subgroup analysis, we found a small correlation ($r = 0.14$, 0.10 to 0.19) between trust and patients' health behaviours and a moderate correlation between trust and health-related subjective experiences ($r = 0.37$, 0.27 to 0.47; [Table 1](#)). Our final stratification of the health-related subjective experiences showed a large association between trust and patient satisfaction ($r = 0.57$, 0.49 to 0.64) and small associations between trust and health-related quality of life ($r = 0.18$, 0.14 to 0.22) and symptom-related outcomes ($r = 0.13$, 0.04 to 0.22). Small to moderate between-study heterogeneity remained unexplained in most explorative analyses ([Table 1](#)).

Publication bias

Despite a significant Egger's regression test in the overall analysis as well as in the analysis using only studies with a health behaviour as outcome ($P < 0.001$), the fail-safe N and the

Table 1. Associations between Trust in the Health Care Professional and Health Outcome stratified according to the Outcome Dimension.

Analysis	N of studies (patients)	r	95% CI	p	τ^2	Publication bias		
						Egger test (p)	Fail-safe N	Trim & Fill test
Overall	47 (34 817)	0.24	0.19, 0.29	<0.001	0.03	<0.001	9 328	0.24
Objective	15 (7 867)	-0.02	-0.08, 0.03	0.430	0.01	0.518	0	-0.02
Observer-rated	2 (706)	0.10	-0.16, 0.36	0.445	0.04	-	-	-
Subjective, self-rated	42 (30 943)	0.30	0.24, 0.35	<0.001	0.04	<0.001	10 532	0.30
Behaviour	21 (26 642)	0.14	0.10, 0.19	<0.001	0.01	0.010	1 857	0.14
Experience	29 (10 229)	0.37	0.27, 0.47	<0.001	0.09	0.226	23 39	0.37
Satisfaction	15 (5 141)	0.57	0.49, 0.64	<0.001	0.04	0.636	8 007	0.57
HRQoL	5 (1 816)	0.18	0.14, 0.22	<0.001	<0.01	0.134	93	0.17
Symptom-related	13 (4 285)	0.13	0.04, 0.22	0.004	0.02	0.333	263	0.13

Note. Study is used as the the unit of analysis. Ns of Subanalysis (objective, observer-rated and subjective, self-rated) do not add up to 34 817, since several studies included more than one outcome. r = correlation; CI = confidence interval; τ^2 = variability between studies; — = no estimate provided due to small number of included studies; HRQoL = health-related quality of life.

doi:10.1371/journal.pone.0170988.t001

Table 2. Meta-Regressions of the Association between Trust in the Health Care Professional and Health Outcome according to Study Quality, Duration of Treatment.

Moderator	N of studies	B	95% CI	p	τ^2
STROBE (study quality)	47	-0.033	-0.07, -0.003	0.074	0.06
Duration of treatment	6	-0.001	-0.01, 0.01	0.751	0.02

Note. N = number of studies included in the analysis; B = unstandardized regression coefficient from meta-regression; CI = confidence interval; τ^2 = variability between studies for the intercept of the model only; STROBE = Strengthening the Reporting of Observational Studies in Epidemiology (study quality).

doi:10.1371/journal.pone.0170988.t002

trim-and-fill analyses indicate a small risk for publication bias (Table 1 and S3 File, which shows the funnel plots).

Meta-regressions and subgroup analyses

Meta-regressions with study quality and treatment duration as predictors showed that the association between trust and outcome depended on study quality, with smaller associations in higher quality studies (Table 2 and S4 File, which shows the scatter plots of the meta-regressions). Furthermore, based on six reports of treatment duration, we found the correlation between trust and outcome to be independent of treatment duration (Table 2 and S4 File). In both meta-regressions one outlier study was identified. Repeated analyses excluding the respective outliers showed similar results as the initial analyses (see S4 File).

Stratified analyses showed some variation in associations between trust and outcome when studies were conducted in different geographical regions but no variation when different trust questionnaires were used (Table 3). Further analyses showed considerable differences when correlational data vs. binary data were reported and finally a meta-analysis restricted to the prospective studies showed a comparable correlation as the analysis including all studies (see Table 3 and S2 File, which shows the respective forest plots).

Table 3. Associations between Trust in the Health Care Professional and Health Outcome stratified according to the Geographic Region, Trust Questionnaire and Type of Data.

Subgroup Analysis	N of studies (patients)	r	95% CI	p	τ^2
Country ^a					
Asia	2 (536)	0.13	0.10, 0.15	<0.001	<0.01
Australia	2 (665)	0.35	-0.31, 0.79	0.298	0.25
Europe	6 (848)	0.36	0.22, 0.48	<0.001	0.03
North America	34 (31 780)	0.22	0.16, 0.28	<0.001	0.04
Trust Questionnaire ^a					
Trust in Physician Scale ^b	24 (17 650)	0.27	0.19, 0.35	<0.001	0.04
Other Trust Questionnaires	23 (17 167)	0.19	0.13, 0.25	<0.001	0.02
Type of Data					
Correlational Data	33 (19492)	0.27	0.26, 0.28	<0.001	<0.01
Binary Data	15 (206867)	0.05	0.03, 0.08	<0.001	0.05
Prospective Data	4 (1584)	0.23	-0.02, 0.45	0.072	0.06

Note. Study is used as the the unit of analyses. r = correlation; CI = confidence interval.

^a Studies do not add up to 47 due to missing information.

^b Anderson LA, Dedrick RF. Development of the Trust in Physician Scale: a measure to assess interpersonal trust in patient-physician relationships. *Psychol Rep* 1990;67:1091–100.

doi:10.1371/journal.pone.0170988.t003

Discussion

We observed a significant association between trust in the health care professional and health outcome. However, results differed with regard to the outcome dimension, with small and non-significant correlations for objective and observer-rated outcomes and a moderate association with self-rated subjective outcomes. The association between trust and outcome was smaller in high quality studies. Interestingly, the observed association between trust and health appeared to be smaller when binary data were reported and larger, when correlations were reported, and smaller in North America and Asia compared to Europe and Australia.

To the best of our knowledge, this is the first meta-analysis that provides an empirical estimate of the association between trust in the health care professional and health outcome. Our analyses included 34 817 participants from 47 studies in different clinical settings—i.e. disorders and treatment duration varied across as well as within studies and studies were conducted in diverse geographic regions with possibly diverse health care systems.

In order to reduce pragmatic heterogeneity, we included only studies that used a validated trust questionnaire, and we clustered health outcomes into different health outcome dimensions. Furthermore, we checked whether the type of questionnaire moderated study results and confirmed the robustness of the overall finding in a subgroup of 24 studies that all used the same trust questionnaire.

Our meta-analysis has several limitations. First, there are indications that our overall results may be overestimated. We found smaller associations in higher quality studies as well as in larger studies and the significant Egger's test indicates a lack of small-scaled studies with non-significant correlations (Table 2 and S3 File for funnel plots). However, the large fail-safe N in our analysis indicated a very low risk for a non-significant overall association between trust in the health care professional and health outcome. Second, we were not able to satisfactorily estimate the impact of potential moderators. Here, a reasonable classification of patient characteristics was not possible due to the fact that most samples were mixed with respect to potentially relevant characteristics (e.g. ethnicity or disorder). Also, only six studies reported details on treatment duration. This resulted in a low power of our meta-regression. Statistical heterogeneity, however, was small to moderate in most analyses. This indicates only a small risk for the presence of strong moderators. Finally and most importantly, our analyses do not allow causal interpretations of the observed association between trust in the health care professional and health outcome, since the vast majority of included studies were cross-sectional. However, a subgroup-analysis including only the four prospective studies that assessed trust before the outcome assessment, showed the same moderately sized association between trust and health outcomes as the overall analysis including prospective and cross-sectional studies.

Patients' trust in the health care professional may best be conceptualized as a contextual factor of treatment effects [10, 11, 34]. Indeed previous work mentions trust in relation to the patient-clinician relationship, which is also embraced by the umbrella-term of contextual factors. Systematic reviews and meta-analyses of RCTs have found small to moderate associations between such factors on health outcomes [13, 34]. Thus, given the risk that the overall association between trust and outcome was overestimated in our meta-analysis, a slightly reduced association in our study could be considered as complementing previous findings.

The differential finding for the type of outcome dimension—i.e. larger associations for the self-rated subjective outcomes and a small or even non-significant association for objective outcomes—has already been described by Beecher in 1955 and has been confirmed for diverse contextual factors ever since [35–38].

A conceptual proximity between trust and subjective outcomes could explain the observed large associations [6, 39]. Indeed, data from a large-scale survey empirically confirms a

meaningful association between trust in the health care professional and subjective health [8]. However, since meta-analyses described an enhanced risk of bias with regard to self-rated subjective outcomes, possibly due to inadequate allocation concealment and inappropriate blinding [40, 41], the particularly large associations in our analyses may result—at least partly—from an upward bias.

The non-significant association between trust and objective health outcomes in our meta-analyses may be seen as confirming the previous findings and thus, reflect a *de facto* absence of such an association. However, we observed a significant correlation between trust and self-rated subjective outcomes, which in turn have been associated with objective outcomes [42–44]. Therefore, it could be argued that a possible association between trust and objective outcomes depends on trust-sensitive subjective variables, such as adherence to medication or patient satisfaction with treatment [45, 46]. We could not test the assumption of such a causal chain, however, since the included studies did not stratify their results with regard to possible trust-sensitive subjective variables. Considering that the establishment of interpersonal trust has been described as evolving continuously rather than being a rigid state [47], the cross-sectional study design of most included studies might have hindered the detection of a possible time-delayed impact of trust on objective outcomes.

We found a larger association in studies that reported correlations and a lower association in studies that reported binary data. This pattern was not due to a predominant use of objective outcomes in the studies with binary data and subjective outcomes in the studies with correlational data: Among the studies that reported correlations 33.33% reported objective data, and among the studies that reported binary data 26.67% reported objective data. With regard to the differences in the observed trust-outcome associations in Europe, North America, Australia, and Asia, which are in part also reflected in a recent survey on trust in physicians [48], it is tempting to assume differences in health care systems [49] or in social and cultural factors [50] to account for this finding. However, the available dataset of our meta-analysis did not allow for the testing of these assumptions.

The summarized data indicate that patients report more beneficial health behaviours, higher satisfaction and health-related quality of life, but also better symptom-oriented subjective outcomes when they had higher trust in their health care professional. These findings substantiate the asserted fundamental role of patients' trust in the context of patient-centered care [9, 18]. It appears tempting to implement the suggested predictors of trust [51, 52] in clinical care as a feasible and possibly cost-effective way to enhance trust and, thus, health outcomes. However, it remains debatable whether such action suffices [18]. First, it must be taken into account that trust and health outcomes may mutually affect each other [6, 43]. Also, trust has been argued to be highly sensitive to more distal, i.e. political, social, and cultural processes [50], which could possibly be reflected in the geographical differences in our analyses. Thus, a sustainable investment should encompass the micro (e.g. patient-provider relationship, quality of health care provision) as well as the macro level (e.g. organisation, financing, and structure of the health care service) [53].

Previous research has proposed different models on how trust could influence health outcomes [45]. For example, Lee and Lin proposed that patients' trust influences the health outcomes via patient disclosure, the placebo effect, compliance, and the physician's caring behaviour [54]. Although this model particularly focuses on the association between patients' trust and health outcomes, it lacks conceptual clarity, for instance with respect to the definition of health outcomes and the influence of more distal factors as well as possible mediators and moderators. Our study highlights the need to differentiate between outcome dimensions. For instance, in accordance with the model proposed by Wampold and Imel, trust in the health care professional might have a positive impact on subjective health (e.g. beneficial health

behaviors), which might then, in turn, lead to improvements on objective outcomes [11]. Accordingly, a conceptual clarification of key variables would not only advance the debate on trust, but also allow the deduction of empirically testable hypotheses. The complex interactions between trust and health outcome, including potentially time-delayed effects, reverse causality, as well as the existence of moderators and mediators should further be investigated in prospective studies. Finally, the influence of distal factors (organizational, political, social as well as cultural variables) on the association between trust and outcome needs to be tested and considered in an all-encompassing model.

Conclusion

Across diverse clinical settings, patients reported to be more satisfied with treatment, to show more beneficial health behaviours, less symptoms and higher quality of life when they had higher trust in their health care professional. But there was no association between trust and observer-rated or trust and objective health outcomes. Although further studies are required to test the direction of the association between trust and health outcome, trust in the health care professional may not only be a deontological constituent of clinical care [55], but it might also be consequential for patients' treatment satisfaction, health behaviours, symptom severity and quality of life.

Supporting information

S1 File. Search Strategies.

(PDF)

S2 File. Forest Plots of all conducted meta-analyses and subgroup analyses.

(PDF)

S3 File. Funnel Plots.

(PDF)

S4 File. Scatter Plots of meta-regressions.

(PDF)

S1 Table. Characteristics of Studies Included in the Meta-Analysis.

(PDF)

Acknowledgments

We thank Paul Bain at Countway Library, Harvard Medical School, for conducting the systematic literature search.

Author contributions

Conceptualization: PK JG JK HG JB.

Data curation: JB.

Formal analysis: JB HG.

Funding acquisition: JB HG.

Investigation: SH CW JB.

Methodology: JB HG.

Project administration: JB HG.

Resources: JG.

Supervision: HG JG.

Visualization: JB HG JG.

Writing – original draft: JB.

Writing – review & editing: HG JG JK.

References

1. Coulter A. Patients' views of the good doctor. *Bmj*. 2002; 325(7366):668–9. PubMed Central PMCID: PMC1124204. PMID: [12351342](#)
2. Mechanic D, Schlesinger M. The impact of managed care on patients' trust in medical care and their physicians. *Jama*. 1996; 275(21):1693–7. PMID: [8637148](#)
3. General Medical Council. Duties of a doctor: The duties of a doctor registered with the General Medical Council [cited 2016 27 Jan]. Available from: http://www.gmc-uk.org/guidance/good_medical_practice/duties_of_a_doctor.asp.
4. Nursing & Midwifery Council (NMC). The Code for nurses and midwives: Read The Code online 2015 [cited 2016 27 Jan]. Available from: <http://www.nmc.org.uk/standards/code/read-the-code-online/>.
5. The British Psychological Society. Code of Ethics and Conduct: Guidance published by the Ethics Committee of the British Psychology Society 2009 [cited 2016 27 Jan]. Available from: http://www.bps.org.uk/system/files/documents/code_of_ethics_and_conduct.pdf.
6. Hall MA, Dugan E, Zheng BY, Mishra AK. Trust in physicians and medical institutions: What is it, can it be measured, and does it matter? *Milbank Q*. 2001; 79(4):613–+. doi: [10.1111/1468-0009.00223](#) PMID: [11789119](#)
7. Calnan M, Rowe R. Researching trust relations in health care: conceptual and methodological challenges—introduction. *Journal of health organization and management*. 2006; 20(5):349–58. doi: [10.1108/14777260610701759](#) PMID: [17087399](#)
8. Croker JE, Swancutt DR, Roberts MJ, Abel GA, Roland M, Campbell JL. Factors affecting patients' trust and confidence in GPs: evidence from the English national GP patient survey. *BMJ open*. 2013; 3(5). PubMed Central PMCID: PMC3657663.
9. Sakallaris BR, Miller WL, Saper R, Kreitzer MJ, Jonas W. Meeting the Challenge of a More Person-centered Future for US Healthcare. *Glob Adv Health Med*. 2016; 5(1):51–60. PubMed Central PMCID: PMC4756778. doi: [10.7453/gahmj.2015.085](#) PMID: [26937314](#)
10. Benedetti F. Placebo and the new physiology of the doctor-patient relationship. *Physiological reviews*. 2013; 93(3):1207–46. doi: [10.1152/physrev.00043.2012](#) PMID: [23899563](#)
11. Wampold BE, Imel ZE. *The great psychotherapy debate: The evidence for what makes psychotherapy work*: Routledge; 2015.
12. Lee YY, Lin JL. How much does trust really matter? A study of the longitudinal effects of trust and decision-making preferences on diabetic patient outcomes. *Patient Educ Couns*. 2011; 85(3):406–12. doi: [10.1016/j.pec.2010.12.005](#) PMID: [21269794](#)
13. Kelley JM, Kraft-Todd G, Schapira L, Kossowsky J, Riess H. The Influence of the Patient-Clinician Relationship on Healthcare Outcomes: A Systematic Review and MetaAnalysis of Randomized Controlled Trials. *Plos One*. 2014; 9(4).
14. Graham JL, Grimes RM, Slomka J, Ross M, Hwang LY, Giordano TP. The Role of Trust in Delayed HIV Diagnosis in a Diverse, Urban Population. *Aids Behav*. 2013; 17(1):266–73. doi: [10.1007/s10461-011-0114-9](#) PMID: [22198316](#)
15. Freburger JK, Callahan LF, Currey SS, Anderson LA. Use of the trust in physician scale in patients with rheumatic disease: Psychometric properties and correlates of trust in the rheumatologist. *Arthritis Rheum-Arthr*. 2003; 49(1):51–8.
16. Boothroyd RA, Rocca TD, Chen HJ. Predictors of enrollees' satisfaction with a county-sponsored indigent health care plan. *Eval Health Prof*. 2008; 31(1):81–103. doi: [10.1177/0163278707311882](#) PMID: [18180517](#)
17. Durant RW, McClure LA, Halanych JH, Lewis CE, Prineas RJ, Glasser SP, et al. Trust in Physicians and Blood Pressure Control in Blacks and Whites Being Treated for Hypertension in the Regards Study. *Ethnic Dis*. 2010; 20(3):282–9.

18. Pearson SD, Raeke LH. Patients' trust in physicians: many theories, few measures, and little data. *Journal of general internal medicine*. 2000; 15(7):509–13. PubMed Central PMCID: PMC1495476. doi: [10.1046/j.1525-1497.2000.11002.x](https://doi.org/10.1046/j.1525-1497.2000.11002.x) PMID: [10940139](https://pubmed.ncbi.nlm.nih.gov/10940139/)
19. Wright EB, Holcombe C, Salmon P. Doctors communication of trust, care, and respect in breast cancer: qualitative study. *Brit Med J*. 2004; 328(7444):864–7. doi: [10.1136/bmj.38046.771308.7C](https://doi.org/10.1136/bmj.38046.771308.7C) PMID: [15054034](https://pubmed.ncbi.nlm.nih.gov/15054034/)
20. von Elm E, Altman DG, Egger M, Pocock SJ, Gotsche PC, Vandenbroucke JP, et al. The Strengthening of Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet*. 2007; 370(9596):1453–7. doi: [10.1016/S0140-6736\(07\)61602-X](https://doi.org/10.1016/S0140-6736(07)61602-X) PMID: [18064739](https://pubmed.ncbi.nlm.nih.gov/18064739/)
21. Cohen J. *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum; 1988.
22. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. *Introduction to Meta-Analysis* Chichester: John Wiley & Sons Ltd; 2009.
23. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *Brmj*. 1997; 315(7109):629–34. PubMed Central PMCID: PMC2127453. PMID: [9310563](https://pubmed.ncbi.nlm.nih.gov/9310563/)
24. Borenstein M. Software for publication bias. In: Rothstein HR, Sutton AJ, Borenstein M, editors. *Publication Bias in Meta-analysis: Prevention, Assessment and Adjustments*. Chichester, England: John Wiley & Sons Ltd; 2005. p. 193–220.
25. Rosenthal R. *Meta-analytic procedures for social research*. Newbury Park, CA: Sage; 1991.
26. Duval S, Tweedie R. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*. 2000; 56(2):455–63. PMID: [10877304](https://pubmed.ncbi.nlm.nih.gov/10877304/)
27. Rucker G, Schwarzer G, Carpenter JR, Schumacher M. Undue reliance on I(2) in assessing heterogeneity may mislead. *BMC medical research methodology*. 2008; 8:79. PubMed Central PMCID: PMC2648991. doi: [10.1186/1471-2288-8-79](https://doi.org/10.1186/1471-2288-8-79) PMID: [19036172](https://pubmed.ncbi.nlm.nih.gov/19036172/)
28. Anderson LA, Dedrick RF. Development of the Trust in Physician scale: a measure to assess interpersonal trust in patient-physician relationships. *Psychological reports*. 1990; 67(3 Pt 2):1091–100.
29. Hall MA, Zheng BY, Dugan E, Camacho F, Kidd KE, Mishra A, et al. Measuring patients' trust in their primary care providers. *Med Care Res Rev*. 2002; 59(3):293–318. PMID: [12205830](https://pubmed.ncbi.nlm.nih.gov/12205830/)
30. Safran DG, Kosinski M, Tarlov AR, Rogers WH, Taira DH, Lieberman N, et al. The Primary Care Assessment Survey: tests of data quality and measurement performance. *Medical care*. 1998; 36(5):728–39. PMID: [9596063](https://pubmed.ncbi.nlm.nih.gov/9596063/)
31. Linden M, Nather J, Wilms U. Definition, Bedeutung und Messung der Krankheitskonzepte von Patienten: Die Krankheitskonzeptskala (KK-Skala) für schizophrenen Patienten. *Neurol Psychiatr* 1988; 56:35–43.
32. Weech-Maldonado R, Carle A, Weidmer B, Hurtado M, Ngo-Metzger Q, Hays RD. The Consumer Assessment of Healthcare Providers and Systems (CAHPS) Cultural Competence (CC) Item Set. *Medical care*. 2012; 50(9):S22–S31.
33. Pfaff H, Freise D, Mager G, Schrappe M. *Der Kölner Patientenfragebogen (KPF): Entwicklung und Validierung eines Fragebogens zur Erfassung der Einbindung des Patienten als Kotherapeuten*. Sankt Augustin: Asgard-Verlag; 2003.
34. Di Blasi Z, Harkness E, Ernst E, Georgiou A, Kleijnen J. Influence of context effects on health outcomes: a systematic review. *Lancet*. 2001; 357(9258):757–62. PMID: [11253970](https://pubmed.ncbi.nlm.nih.gov/11253970/)
35. Beecher HK. The Powerful Placebo. *Jama-J Am Med Assoc*. 1955; 159(17):1602–6.
36. Hrobjartsson A, Gotzsche PC. Placebo interventions for all clinical conditions. *Cochrane Db Syst Rev*. 2010;(1).
37. Schwarz KA, Buchel C. Cognition and the Placebo Effect—Dissociating Subjective Perception and Actual Performance. *Plos One*. 2015; 10(7).
38. Wechsler ME, Kelley JM, Boyd IOE, Dutile S, Marigowda G, Kirsch I, et al. Active Albuterol or Placebo, Sham Acupuncture, or No Intervention in Asthma. *New Engl J Med*. 2011; 365(2):119–26. doi: [10.1056/NEJMoa1103319](https://doi.org/10.1056/NEJMoa1103319) PMID: [21751905](https://pubmed.ncbi.nlm.nih.gov/21751905/)
39. Thom DH, Ribisl KM, Stewart AL, Luke DA, Phys STS. Further validation and reliability testing of the Trust in Physician Scale. *Medical care*. 1999; 37(5):510–7.
40. Moustgaard H, Bello S, Miller FG, Hrobjartsson A. Subjective and objective outcomes in randomized clinical trials: definitions differed in methods publications and were often absent from trial reports. *J Clin Epidemiol*. 2014; 67(12):1327–34. doi: [10.1016/j.jclinepi.2014.06.020](https://doi.org/10.1016/j.jclinepi.2014.06.020) PMID: [25263546](https://pubmed.ncbi.nlm.nih.gov/25263546/)
41. Wood L, Egger M, Gluud LL, Schulz KF, Juni P, Altman DG, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *Brit Med J*. 2008; 336(7644):601–5. doi: [10.1136/bmj.39465.451748.AD](https://doi.org/10.1136/bmj.39465.451748.AD) PMID: [18316340](https://pubmed.ncbi.nlm.nih.gov/18316340/)

42. Detmar SB, Muller MJ, Schornagel JH, Wever LDV, Aaronson NK. Health-related quality-of-life assessments and patient-physician communication—A randomized controlled trial. *Jama-J Am Med Assoc.* 2002; 288(23):3027–34.
43. Doyle C, Lennox L, Bell D. A systematic review of evidence on the links between patient experience and clinical safety and effectiveness. *BMJ open.* 2013;3(1).
44. Simpson SH, Eurich DT, Majumdar SR, Padwal RS, Tsuyuki RT, Varney J, et al. A meta-analysis of the association between adherence to drug therapy and mortality. *Brit Med J.* 2006; 333(7557):15–8. doi: [10.1136/bmj.38875.675486.55](https://doi.org/10.1136/bmj.38875.675486.55) PMID: [16790458](https://pubmed.ncbi.nlm.nih.gov/16790458/)
45. Giordano GN, Lindstrom M. Trust and health: testing the reverse causality hypothesis. *J Epidemiol Commun H.* 2016; 70(1):10–6.
46. Street RL, Makoul G, Arora NK, Epstein RM. How does communication heal? Pathways linking clinician–patient communication to health outcomes. *Patient Educ Couns.* 2009; 74(3):295–301. doi: [10.1016/j.pec.2008.11.015](https://doi.org/10.1016/j.pec.2008.11.015) PMID: [19150199](https://pubmed.ncbi.nlm.nih.gov/19150199/)
47. Lewicki RJ, Tomlinson EC, Gillespie N. Models of interpersonal trust development: Theoretical approaches, empirical evidence, and future directions. *J Manage.* 2006; 32(6):991–1022.
48. Blendon RJ, Benson JM, Hero JO. Public trust in physicians—U.S. medicine in international perspective. *The New England journal of medicine.* 2014; 371(17):1570–2. doi: [10.1056/NEJMp1407373](https://doi.org/10.1056/NEJMp1407373) PMID: [25337746](https://pubmed.ncbi.nlm.nih.gov/25337746/)
49. Shortell SM, Waters TM, Clarke KWB, Budetti PP. Physicians as double agents—Maintaining trust in an era of multiple accountabilities. *Jama-J Am Med Assoc.* 1998; 280(12):1102–8.
50. Gopichandran V, Wouters E, Chetlapalli SK. Development and validation of a socioculturally competent trust in physician scale for a developing country setting. *BMJ open.* 2015; 5(4).
51. Keating NL, Gandhi TK, Orav EJ, Bates DW, Ayanian JZ. Patient characteristics and experiences associated with trust in specialist physicians. *Arch Intern Med.* 2004; 164(9):1015–20. doi: [10.1001/archinte.164.9.1015](https://doi.org/10.1001/archinte.164.9.1015) PMID: [15136312](https://pubmed.ncbi.nlm.nih.gov/15136312/)
52. Ommen O, Thuem S, Pfaff H, Janssen C. The relationship between social support, shared decision-making and patient's trust in doctors: a cross-sectional survey of 2,197 inpatients using the Cologne Patient Questionnaire. *International journal of public health.* 2011; 56(3):319–27. doi: [10.1007/s00038-010-0212-x](https://doi.org/10.1007/s00038-010-0212-x) PMID: [21076932](https://pubmed.ncbi.nlm.nih.gov/21076932/)
53. Gilson L. Trust and the development of health care as a social institution. *Soc Sci Med.* 2003; 56(7):1453–68. PMID: [12614697](https://pubmed.ncbi.nlm.nih.gov/12614697/)
54. Lee Y-Y, Lin JL. Linking patients' trust in physicians to health outcomes. *British Journal of Hospital Medicine (17508460).* 2008; 69(1).
55. Charon R. The patient-physician relationship. *Narrative medicine: a model for empathy, reflection, profession, and trust.* *Jama.* 2001; 286(15):1897–902. PMID: [11597295](https://pubmed.ncbi.nlm.nih.gov/11597295/)

Experimental manipulation of interpersonal trust: an adaption of the trust game

Johanna Birkhäuser^{1*}, MSc; Christoph Werner¹, BSc; Jens Gaab¹, PhD

¹Department of Psychology, University of Basel, Basel, Switzerland

*Corresponding author:

Email: j.birkhaeuer@unibas.ch

Abstract

Trust is considered a necessary prerequisite of consequential interpersonal encounters, such as societal, political, economical and therapeutic exchange relationships. With regard to the wealth of empirical evidence on trust, it needs to be noted that while there is a surplus of studies investigating trust as a dependent, covarying or predicting variable, the experimental manipulation of trust has received considerable less attention. Here, we examined and adapted the trust game, an established paradigm in economic psychology, and assessed the effects of random allocation to either a trustful or a non-trustful interaction on subjective as well as behavioral trust measures. The results show that trust can be experimentally manipulated.

Introduction

Interpersonal trust is the fundament for societal [1], organizational [2], family [3], political [4, 5] and economical exchange relationships [6, 7]. Furthermore interpersonal trust is also considered fundamental for the clinical encounter, although its specific role in the treatment process still needs to be elucidated [8, 9]. In the clinical setting, a variety of observational studies have investigated the relationship between trust and health outcome [10]. While these studies provide correlational data, they cannot give evidence for causality. Also, some trials evaluated so-called trust trainings, which set out to improve patients' trust in their doctors, although with sobering effects [11]. On the basis of these findings, it has been noted “that trust is important for its potential therapeutic effects although evidence to support such claims is still in short to supply mainly because of the lack of intervention studies or quasi-experimental studies examining the effect of trust on health outcomes...” [8]. This lack might be due to the fact that trust is a multifactorial and complex concept that differs among situations and changes over time [12].

In the domain of economic psychology, the trust game by Berg, Dickhaut and McCabe (1995) is a well-elaborated and replicated method to measure trust in an experimental setting, including the sequential momentary exchange of two parties, with no predefined contract or agreement [13]. Here, the vast majority of trust game studies focus on trust as a dependent variable [14-16], although there are noteworthy exceptions, with experimental induction of distrust in two studies so far [17, 18]. However, these (dis-)trust manipulations were uncontrolled for a trust condition.

Furthermore, whereas the trust game has predominantly been seen in the context of expected maximization of the monetary payoffs, several authors have postulated that the participants' behavior is better understood in the framing of a social interaction [19]. Noteworthy and in consequence, Castelfranchi & Falcone (2000) highlighted the limits of the

strategic tradition of trust for the transfer to real life situations, assuming that the trustor forms a theory of mind of the trustee, including personality, morality, and shared values, not realizable by the use of a computerized partner [20]. Since trust-relevant situations are complex and interactive, it is both difficult to transfer results from an anonymous trust game as much as to exclude or control for confounding variables in a natural environment, e.g. in a clinical setting [21, 22]. This is especially true with regard to anonymity, a factor extensively debated in trust game research. It has been hypothesized that when anonymity is disregarded in the trust game other factors might play a substantial role as the protection of the own reputation or the fear of retribution [13].

Considering on the one hand the assumed importance of trust for various domains and on the other hand the lack of an established paradigm to experimentally manipulated trust, the primary objective of the current study was to investigate whether the trust game might provide a tool to induce trust by using two fixed conditions (a trust and a non trust condition). We tested this by assigning the role of the trustee to a non-anonymous experimenter, thereby providing a more naturalistic setting, in order to compare participants' subjective experience with their behavioral response, i.e. transferred money.

Methods

Participants and procedure

In total, 106 male participants were enrolled in the current study (mean age = 24.5 years, SD = 4.7 years). Participants were recruited by advertisement at the University of Basel, via two online recruitment systems of the University of Basel (BAPS-Sona, <http://baps.sona-systems.com> and markt.unibas.ch). Potential participants were previously checked for inclusion by the use of an Internet survey. Inclusion criteria were male gender, age between 18 to 40 years and sufficient German language skills. Exclusion criteria were any acute or chronic disease by self-report.

Trust game

The study was conducted at the Department of Psychology of the University of Basel. The experiment consisted of one session (2 hours). The participants played a trust game, adapted from Berg, Dickhaut and McCabe (1995), but were told they were going to play a cognitive task in its context they could win some additional money [23]. Some changes were made in order to manipulate the participant's level of trust. Starting with an endowment of 1 start unit in each round, the participant could decide to send a particular portion of their endowment to the investigator. In total, five rounds were played. In round 1 and 2, the participant could send 25%, 50% or 75% of their endowment; while in round 3, 4 and 5 also 0% or 100% could be sent. After the participant sent the money, the amount was tripled. The investigator (in the role of the trustee) in return passed a predefined amount of money back, depending on the condition (see below).

Participants were randomly assigned (by lot) to the non-trust or trust condition immediately before the trust game started. In the non-trust condition the investigator returned

50%, 0%, 25%, 0% and 0% of the amount sent (round 1 to 5, respectively). The sent amount was tripled. With regard to the trust condition, based on the fairness heuristic that trust should be facilitated when the decision yields maximum rewards for both parties, the experimenter always returned 50% of the tripled amount in order to create a fair exchange [19]. The amount passed by the sender was defined to capture the behavioral response of trust [23]. The gain was disbursed at the end of the experiment and no feedback was given from the experimenter.

Before and after the trust game we employed the State-Trait-Anxiety-Inventory (STAI) to measure anxiety. The STAI is a standardized self-report questionnaire, measuring anxiety both as state (X1) as well as a trait (X2). Both scales contain 20 items and a four-point ranking scale, ranging from 1 = almost never to 4 = almost always. Cronbach's alpha for both scales is found to be $\alpha = .90$ [24]. Negative items were recoded. Higher items indicate a higher level of anxiety. Also, we used the State-Trait-Anger-Expression-Inventory (STAXI) before and after the trust game, which assesses five subscales of anger: state anger (10 items), trait anger (10 items), and anger expression (24 items). Items are rated on a four-point ranking scale from 1 = not at all to 4 = very much. Cronbach's alpha coefficient ranges from .75 to .82 for the different scales [25]. As a subjective measure of trust, participants were asked to indicate how trustful they perceived the investigator on a 10 cm adapted visual analogue scale (VAS) before (baseline) and after the trust game [26]. The intensity scale ranged from 0 (no trust) to 100 (high trust). Higher scores indicated higher trust. All instructions given to participants as well as the experiment's technical procedure followed a standardized experimental protocol. The study was approved by the 'Ethikkommission Nordwest- und Zentralschweiz (EKNZ)' and were registered by the 'Clinical Trial Registration' (EKNZ 2014-396, URL:<https://clinicaltrials.gov/ct2/results?term=EKNZ+2014-396&Search=Search>). For compensation, participants received either study credits or CHF 50.

Statistical analysis

To test whether the trust condition was associated with differences in the trust rating of the experimenter, we conducted an analysis of covariance (ANCOVA), including trust-condition as the fixed factor and the baseline trust measurement as the covariate. In a second step, we also held STAI- and STAXI pre-post difference constant in order to control for anger or anxiety. With regard to the behavioral level, we investigated whether the trust condition was also associated with differences in the amount send by the participant. Therefore we calculated separate chi-square tests as an explanatory analysis. Also separated t-tests were conducted in order to test for differences in mean percentage of the passed amount. To verify the normality, linearity and homoscedasticity assumption of the residuals, we inspected scatter-, box- and QQ-plots and checked for outliers (leverage, Cook's distance and Mahalanobis distance). Since distribution was skewed we excluded one participant (with regard to his values with regard to leverage, Cook's distance and Mahalanobis distance). For all tests the significance level was set to .05. Statistical analyses were performed using IBM SPSS Statistics Version 22.

Results

Subjective trust rating: Visual analogue scale on trust (VAS)

Overall, 106 participants underwent the experimental manipulation of trust (trust condition $N = 53$ distrust condition $N = 53$). Participants in the two conditions did not differ regarding mean age in years, baseline scores in the psychometric assessment of anxiety and anger as well as in baseline trust-rating (see Table 1). The experimental variation of trust led to significant time by group interaction effect for anger ($F(1, 104) = 7.24, p = .008, \eta_p^2 = .065$) but not for anxiety ($F(1, 102) = .078, \eta_p^2 = .030$). Controlling for the baseline subjective trust-rating as covariate, the post ratings of subjective trust rating showed a main effect of the trust condition after the trust game ($F(1, 102) = 4.47, p = .037, \eta_p^2 = .042$), with participants in the trust condition rating the experimenter as significantly more trustworthy. For exploratory analysis, a mixed model ANOVA showed a significant interaction ($F(1, 104) = 5.08, p = .026, \eta_p^2 = .047$), which reflects a decrease in the non-trust condition after the trust game, whereas for the trust condition no difference in the subjective trust rating after the trust game was found (Fig 1). Controlling for pre-post differences of anxiety and anger, the previously observed significant effects were attenuated, yet still observable with a alpha error below 10% ($F(1, 100) = 3.78, p = .055, \eta_p^2 = .036$) anger ($F(1, 102) = 2.88, p = .093, \eta_p^2 = .027$ (see Table 2).

Fig 1. Subjective trust ratings before and after the trust game.

Table 1. Mean and standard deviation of the variables age, STAI, STAXI, trust-rating.

Dependent variable	Non-trust		Trust		P-Value
	Mean	SD	Mean	SD	
Age	24.51	5.15	24.53	4.15	0.938
STAI					
Pre	33.21	7.36	33.23	6.71	0.987
Post	33.63	8.23	32.30	6.63	0.363
STAXI					
Pre	11.30	3.18	10.94	1.89	0.482
Post	12.04	3.16	10.60	93	0.002
Trust VAS (trust rating)					
Pre	88.00	10.57	86.49	14.61	0.544
Post	83.42	16.98	86.28	14.23	0.348

Note. STAI = State-Trait-Anxiety-Inventory, STAXI = State-Trait-Anger-Expression-Inventory, VAS = visual analogue scale, SD = standard deviation.

Table 2. Effects of the trust manipulation for the trust rating and behavioral response with and without controlling for a covariate (anxiety and anger).

Dependent variable	Df	F/t	P-value
Trust rating ^a	1	5.08	0.026
Investment ^{bc}	104	-2.62	0.027
ANCOVA: anxiety as covariate			
Trust rating ^a	1	3.76	0.055
Investment ^b	1	2.04	0.156
ANCOVA: anger as covariate			
Trust rating ^a	1	2.89	0.093
Investment ^b	1	.54	0.466

^aInteraction term: Time point (before vs after the trust game) x trust condition (non-trust vs trust group) with the trust rating as the dependent variable.

^bInvestment for the third round in the trust game.

^cFor this analysis an independent t-test (non-trust vs trust condition) was used.

Behavioral response to the trust manipulation: Investment in trust game

To operationalize behavioral consequences of the experimental trust manipulation, we assessed the investment of participants. In accordance with Haselhuhn (2010) [17], we dichotomized the investment into whether the participants sent money or not and conducted separate chi square tests. In both conditions participants invested on average more than 50% in each round. At round three, i.e. after trustors' behavior was experimentally manipulated and participants responded to the manipulation, a significant difference between non-trust and trust condition was detectable ($\chi^2(1, N = 105) = 5.13, p = .020$). Also, the mean percentage of the money passed by participant after trustors' behavior changed between groups was significantly higher in the trust condition (Fig 2) for round three ($t(77.12) = -2.62, p = .027, d = .439$). For round four and five no difference could be detected between non-trust and trust (round four: $\chi^2(1, N = 105) = 0.29, p > .250$, round five: $\chi^2(1, N = 105) = 0.10, p > .250$ and mean percentage: round four $t(104) = .57, p > .250$ round five $t(104) = -.28, p > .250$). When controlling for anger and anxiety the found difference between trust conditions in round three were no longer significant (anger: $(F(1, 102) = 0.54, p = .466)$; anxiety: $(F(1, 100) = 2.04, p = .156)$ (see Table 2).

Correlation analysis revealed a significant correlation between subjective trust rating (post measure) and investment in round three ($r = .22, p = .023$) and a non-significant correlation for round four and five (round four: $r = -.079, p = .422$, round five: $r = .144, p = .142$).

Fig 2. Investment in the non-trust and trust condition. Percentages beneath the round describe the response percentage sent by the trustee to the participant. Participants can react to the percentage of the trustee one round later.

Discussion

The aim of the present study was to examine the effects of an experimental trust manipulation. We found a significant effect of the trust manipulation on the subjective as well as the behavioral trust level; whereas in the trust condition subjective rating of trust in the experimenter remained the same before and after the trust game, in the non-trust condition subjective trust ratings decreased significantly. Also, after the trust manipulation a significant difference between trust conditions could be detected on the behavioral level, with decreases in the percentage of participants choosing to pass any amount as well as in the mean percentage of passed money. However, these effects of the trust manipulation were transient, as for the two last rounds no differences were found in these behavioral parameters. Also, the trust manipulation led to significant increases in anger (but not anxiety) and changes in these affective parameters influenced the interventions effects on trust measures.

In the following, our results will be discussed in the context of other findings and studies. First, with the employed paradigm, we were able to systematically manipulate the trust level of the participants. Previously, attempts have been made to vary the trust level in the trust game [15, 17, 19]. However, Haselhuhn included a distrust condition without controlling for a trust condition and Burks (2003) as well as Harth and Regner (2016) manipulated the trust condition only in an indirect way. Burks (2003) manipulated the trust level by letting participants play both roles of the trustor and the trustee in the trust game and found that playing both roles reduced the trust level in the participant [19]. Harth and Regner (2016) varied whether participants received feedback about the trustee's amount sent back [15]. Receiving feedback increased the anger level in the participant, which in turn led to a spiral of distrust. However, both studies did not systematically create a trust and a non-trust condition, but rather changed indicators that indirectly influence the trust level, as feedback or changing roles in the game. Our finding corroborates the induction of transient distrust

reported by Haselhuhn et al. (2010) [17]. Furthermore, we were to extend and strengthen these findings through the inclusion of a control group.

Second, despite the observed significant drop in both subjective as well as in behavioral trust measures after the trust manipulation, our results also indicate that a certain amount of trust resists even in the presence of unexpected behavior of the trustor, which comes close to betrayal. This resistance was apparent for both - the subjective experience of trust as well as for the behavioral level: Participants investigated more than 50% in each round, regardless of their condition. Our findings are in line with results from Burnham and colleagues (2000) that participants still sent money in the case of increasing betrayal as even in the distrust condition, participants sent more than at least 19% of their endowment [27].

Third, the subjective experience of trust did not correspond with the behavioral response in rounds four and five, where no differences between trust and non-trust condition with regard to investment were found. Also, there was no correlation between trust rating and investment for round four and five. This seeming dissociation between subjective and behavioral trust operationalization confirms earlier and comparable findings. For example, Glaeser and colleagues (2000) reported that behavior in the trust game was predicted by past trusting behavior but not by the responses of the participant on a – albeit attitudinal – trust survey [28]. With regard to our results it is also possible that a similar decline of subjective trust has occurred, but was not detectable simply because we did not assess this parameter repeatedly, but measured only before and after the trust game. Also, the lack of a correlation might be due to a truly existing dissociation between trust perception and behavior in the fourth and fifth round. Based on our results it cannot be excluded that the subjective experience of trust might deviate from the replying investing behavior of the participant. Outside the frame of the trust game, trusting behavior has been linked to the expectations of

reciprocity rather than perceived trust [26]. Therefore, caution is warranted when equalizing the investment (trust behavior) as a measure of trust experience [13].

Finally, it has been argued that in a non-anonymous trust game participants might be more motivated by fear than by trust [29]. In our non-anonymous experiment, the effect of the trust manipulation was indeed influenced by anger and anxiety of the participants. These findings are in line with previous reports by Capra and colleagues who found that the perception of the trustors payoff (as high or low) depends on the mood state of the participant. Capra highlights that in a strategic interaction like the trust game, identical logical structures (the same payoff) can be perceived differently depending on the participant's mood [30]. With our results we could support this finding, showing that participants did perceive the differences in payoff differently when the mood was held constant (differences between group attenuated). Moreover, Eberl and colleagues describe with regard to the trust game that participants try to test the trustors conscience, checking for a sufficient emotional bonding [31]. In this context temporary emotions (as anger and anxiety state) might form the basis for trust development, which has been ignored largely in trust game research. Putting our results in the context of emotional bonding we suggest, that the trust game and the development of trust needs to be investigated in a more realistic social context, where the role of emotions and relationship towards a second person can be studied in a more systematic way.

To the best of our knowledge this is this first study using the trust game as an independent variable, i.e. systematically manipulating the trust level by changing trustors behavior. However, some limitations must be noted. First, as mentioned before, we only measured the subjective experience of trust before and after the trust game, eliminating the possibility to capture changes across different rounds. This was due to the fact that we try to prevent priming the participants to the topic of trust, thereby introducing a bias to our assessment [32]. Second, it might be reasoned that our trust manipulation is of circular nature,

including a dependent and an independent variable in the same paradigm. However, this problem is inherently embedded in the trust game and we tried to counteract this by including the subjective trust rating as a further variable. Moreover, we cannot fully exclude that our results were confounded by the induction of anger or anxiety. Therefore, we suggest that further research is needed in order to specify the role of emotions in the trust game.

Conclusion

Considering the ubiquitous importance of trust and the lack of experimental studies on this construct, we suggest that the present paradigm provides a possible approach to further elaborate the role of trust in different settings. For example, our paradigm could prove to be useful in the examination of trust in the clinical context. This implementation might also elucidate how trust is build and influences the participants' response behavior in a real life situation, including emotions and the social context of the participant and the trustor. However, this is of course a matter of further studies.

References

1. Luhmann N. Familiarity, confidence, trust: Problems and alternatives. *Trust: Making and breaking cooperative relations*. 2000;6:94-107.
2. Tan HH, Lim AK. Trust in coworkers and trust in organizations. *The Journal of Psychology*. 2009;143(1):45-66.
3. De Carlo I, Widmer ED. The fabric of trust in families: inherited or achieved? *Families and Kinship in Contemporary Europe*: Springer; 2011. p. 216-33.
4. Newton K. Trust, social capital, civil society, and democracy. *International Political Science Review*. 2001;22(2):201-14.
5. Putnam RD. *Bowling alone: America's declining social capital*. Culture and Politics: Springer; 2000. p. 223-34.
6. Fehr E, List JA. The hidden costs and returns of incentives—trust and trustworthiness among CEOs. *Journal of the European Economic Association*. 2004;2(5):743-71.
7. Knack S, Keefer P. Does social capital have an economic payoff? A cross-country investigation. *The Quarterly journal of economics*. 1997:1251-88.
8. Brennan N, Barnes R, Calnan M, Corrigan O, Dieppe P, Entwistle V. Trust in the health-care provider-patient relationship: a systematic mapping review of the evidence base. *Int J Qual Health Care*. 2013;25(6):682-8. doi: 10.1093/intqhc/mzt063. PubMed PMID: 24068242.
9. Pearson SD, Raeke LH. Patients' trust in physicians: many theories, few measures, and little data. *Journal of general internal medicine*. 2000;15(7):509-13.
10. Lee YY, Lin JL. The effects of trust in physician on self-efficacy, adherence and diabetes outcomes. *Soc Sci Med*. 2009;68(6):1060-8. doi: 10.1016/j.socscimed.2008.12.033. PubMed PMID: WOS:000265003400011.

11. McKinstry B, Ashcroft R, Car J, Freeman GK, Sheikh A. Interventions for improving patients' trust in doctors and groups of doctors. *Cochrane Database Syst Rev.* 2006;3.
12. Simpson JA. Psychological foundations of trust. *Curr Dir Psychol Sci.* 2007;16(5):264-8. doi: DOI 10.1111/j.1467-8721.2007.00517.x. PubMed PMID: WOS:000249747900007.
13. Johnson ND, Mislin AA. Trust games: A meta-analysis. *J Econ Psychol.* 2011;32(5):865-89. doi: 10.1016/j.joep.2011.05.007. PubMed PMID: WOS:000295654800023.
14. Boksem MA, Mehta PH, Van den Bergh B, van Son V, Trautmann ST, Roelofs K, et al. Testosterone inhibits trust but promotes reciprocity. *Psychol Sci.* 2013;24(11):2306-14. doi: 10.1177/0956797613495063. PubMed PMID: 24071565.
15. Harth NS, Regner T. The spiral of distrust: (Non-)cooperation in a repeated trust game is predicted by anger and individual differences in negative reciprocity orientation. *Int J Psychol.* 2016. doi: 10.1002/ijop.12257. PubMed PMID: 26865362.
16. Kosfeld M, Heinrichs M, Zak PJ, Fischbacher U, Fehr E. Oxytocin increases trust in humans. *Nature.* 2005;435(7042):673-6. doi: 10.1038/nature03701. PubMed PMID: 15931222.
17. Haselhuhn MP, Schweitzer ME, Wood AM. How Implicit Beliefs Influence Trust Recovery. *Psychol Sci.* 2010;21(5):645-8. doi: 10.1177/0956797610367752. PubMed PMID: WOS:000277837700006.
18. Schweitzer ME, Hershey JC, Bradlow ET. Promises and lies: Restoring violated trust. *Organizational behavior and human decision processes.* 2006;101(1):1-19.
19. Burks SV, Carpenter JP, Verhoogen E. Playing both roles in the trust game. *J Econ Behav Organ.* 2003;51(2):195-216. doi: Pii S0167-2681(02)00093-8 Doi 10.1016/S0167-2681(02)00093-8. PubMed PMID: WOS:000183085400003.

20. Castelfranchi C, Falcone R, editors. Trust is much more than subjective probability: Mental components and sources of trust. *System Sciences, 2000 Proceedings of the 33rd Annual Hawaii International Conference on*; 2000: IEEE.
21. Croker JE, Swancutt DR, Roberts MJ, Abel GA, Roland M, Campbell JL. Factors affecting patients' trust and confidence in GPs: evidence from the English national GP patient survey. *Bmj Open*. 2013;3(5). doi: UNSP e00276210.1136/bmjopen-2013-002762. PubMed PMID: WOS:000330538300075.
22. Wright EB, Holcombe C, Salmon P. Doctors' communication of trust, care, and respect in breast cancer: qualitative study. *Bmj*. 2004;328(7444):864.
23. Berg J, Dickhaut J, McCabe K. Trust, Reciprocity, and Social-History. *Game Econ Behav*. 1995;10(1):122-42. doi: DOI 10.1006/game.1995.1027. PubMed PMID: WOS:A1995RM50400006.
24. Laux L. *Das State-Trait-Angstinventar (STAI): Theoretische Grundlagen und Handanweisung*. 1981.
25. Spielberger CD, Reheiser EC. Assessment of Emotions: Anxiety, Anger, Depression, and Curiosity. *Appl Psychol-Hlth We*. 2009;1(3):271-302. doi: 10.1111/j.1758-0854.2009.01017.x. PubMed PMID: WOS:000208580200001.
26. Tanis M, Postmes T. A social identity approach to trust: Interpersonal perception, group membership and trusting behaviour. *European Journal of Social Psychology*. 2005;35(3):413-24.
27. Burnham T, McCabe K, Smith VL. Friend-or-foe intentionality priming in an extensive form trust game. *J Econ Behav Organ*. 2000;43(1):57-73. doi: Doi 10.1016/S0167-2681(00)00108-6. PubMed PMID: WOS:000089081400003.
28. Glaeser EL, Laibson DI, Scheinkman JA, Soutter CL. Measuring trust. *Q J Econ*. 2000;115(3):811-46. doi: Doi 10.1162/003355300554926. PubMed PMID: WOS:000088808900004.

29. Bies RJ, Tripp TM. Beyond distrust: " Getting even" and the need for revenge. 1996.
30. Capra CM. Mood-driven behavior in strategic interactions. *The American Economic Review*. 2004;94(2):367-72.
31. Eberl P. The development of trust and implications for organizational design: A game- and attribution-theoretical framework. *Schmalenbach Business Review*. 2004;56.
32. Légal J-B, Chappé J, Coiffard V, Villard-Forest A. Don't you know that you want to trust me? Subliminal goal priming and persuasion. *Journal of Experimental Social Psychology*. 2012;48(1):358-60.

EMPLOYING OPEN/HIDDEN ADMINISTRATION IN PSYCHOTHERAPY RE- SEARCH: A RANDOMIZED-CONTROLLED TRIAL OF EXPRESSIVE WRITING

Theresa Tondorf^{o*}, Lisa-Katrin Kaufmann^{*}, Alexander Degel^{*}, Cosima Locher^o, Johanna Birkhäuser^o,
Heike Gerger^o, Ulrike Ehlert^{*}, Jens Gaab^o

Affiliations

^oClinical Psychology and Psychotherapy, Department of Psychology, University of Basel, Switzerland

^{*}Clinical Psychology and Psychotherapy, Institute of Psychology, University of Zurich, Switzerland

Corresponding author

Prof. Dr. Jens Gaab, Clinical Psychology and Psychotherapy, Department of Psychology, University
of Basel, Switzerland, jens.gaab@unibas.ch

Keywords

Expressive writing, Specific effects, Placebo control, Open/Hidden administration

Registry/registration number

German Clinical Trials Register/DRKS00009428

ABSTRACT (word count 149)

Psychotherapy has been shown to be effective, but efforts to prove specific effects by placebo-controlled trials have been practically and conceptually hampered. We propose that adopting open/hidden designs from placebo research would offer a possible way to establish specificity in psychotherapy. Therefore, we tested the effects of providing opposing treatment rationales in an online expressive writing intervention on affect in healthy subjects. Results indicate that it was possible to conduct the expressive writing intervention both covertly and openly, but that participants in the hidden administration condition did not fully benefit from the otherwise effective expressive writing intervention in the long-run. Effect sizes between open and hidden administration groups were comparable to pre-post effect sizes of the intervention. While this finding is important for the understanding of psychotherapy's effects per se, it also proves that alternative research approaches to establish specificity are feasible and informative in psychotherapy research.

INTRODUCTION (word count 739)

There is consensus that psychotherapy is effective, yet identification of accountable factors or processes is subject to controversy (Kazdin, 2007). Whereas in biomedical research the position that "preservation of sound judgment both in the laboratory and in the clinic requires the use of the "double blind" technique" (Beecher, 1955, p. 1606) is considered gold standard, this approach is not fully applicable in psychotherapy research. Although it needs to be noted that even in biomedical research not everything that glitters is gold (Fergusson, Glass, Waring, & Shapiro, 2004), the principle of indistinguishability between comparators is fundamentally hampered in psychotherapy trials both practically and conceptually (Kirsch, 2005). So far, two main approaches have been employed to handle this problem and to establish specificity in psychotherapy research, both with problematic consequences in their own right (Wampold & Imel, 2015, p. 213ff).

First, since it is old lore that "to show that a specific form of psychotherapy (...) produces results not attributable to the non-specific placebo effect it is not sufficient to compare its results with (...) no treatment" (Rosenthal & Frank, 1956, p. 300), psychotherapy is to be compared to psychological placebo to establish specificity of the treatment (Chambless & Hollon, 1998). This seemingly adopts the gold standard of biomedical research and thus is considered good enough to be gentled as "Research-supported psychological treatment" (American Psychological Association, Division 12). However, there is evidence that differences between psychotherapies and psychological placebo conditions¹ are a function of the credibility of the placebo control (Kazdin & Wilcoxon, 1976; Locher et al., 2016) as well as the extent of structural equivalence between comparators (Baskin, Tierney, Minami, & Wampold, 2003) and that controlling for researchers' allegiance towards a particular intervention abolshes observed differences (e.g. Cuijpers et al., 2012). Second, specificity has been sought through direct comparisons of psychotherapies, which usually differ in their characteristic, but not in their incidental treatment constituents (Grünbaum, 1981) and through dismantling or additive component studies, in which presumable characteristic treatment constituents are either omitted or added. However, meta-analyses on respective trials do not substantiate the existence of statistically and/or clinically

¹ Placebo conditions often run under various denominators, such as supportive therapy, nondirective therapy, common factor control, credible attention placebo, and modest contact, but should be and are considered as placebo as they are lacking treatment constituents deemed specific

significant differential effects per se (Ahn & Wampold, 2001; Bell, Marcus, & Goodlad, 2013; Fluckiger, Del Re, & Wampold, 2015; Marcus, O'Connell, Norris, & Sawaqdeh, 2014; Wampold et al., 1997), besides those are related to researchers' allegiance (Munder, Fluckiger, Gerger, Wampold, & Barth, 2012).

In conclusion, the utility of randomized placebo-controlled trials in clinical research "to sort therapeutic wheat from chaff" (Jones & Podolsky, 2015, p. 1502) is not unequivocally warranted in psychotherapy. Given that placebo conditions in psychotherapy cannot be blinded, the main principle of randomized placebo-controlled trials, i.e. to employ different interventions with the same rationale – or put otherwise: to control the incidental, while manipulating the characteristic treatment constituents – is not applicable. However, the strive for specificity – understood here as the identification of characteristic treatment constituents outperforming incidental ones – is important regardless of both the understanding of psychotherapy or of the respective characteristic treatment constituent under investigation.

Here, placebo research designs could come to aid. Using a topsy-turvy approach with regard to randomized placebo-controlled trials, the aim is to control for characteristic, while manipulating incidental treatment constituents. A stringent implementation of this approach is the open versus hidden design, where an otherwise identical intervention is given to subjects either knowing or not knowing when or if the drug will be administered or withdrawn. Differences between conditions are then understood as placebo-like effects, since the intervention itself is not a placebo (e.g. Lund, Vase, Petersen, Jensen, & Finnerup, 2014). This trial design has the merits of being able to specify effects of assumingly incidental treatment constituents, e.g. therapeutic meaning (Moerman & Jonas, 2002), from those of characteristic treatment constituents without the use of a placebo condition and thus has successfully been used in non-pharmacological interventions (e.g. Crum & Langer, 2007; Desharnais, Jobin, Cote, Levesque, & Godin, 1993). However, it has not been tested in the realm of psychotherapy interventions. Employing an online expressive writing intervention, we tested the feasibility of an open/hidden paradigm in psychotherapy and whether the effects of the assumingly characteristic treatment constituent of expressing emotions about a traumatic experience are influenced by the as-

sumingly incidental treatment constituent of providing a therapeutic rationale, i.e. a meaning (Frank, 1986).

METHOD

Design

We conducted a randomized controlled trial with three experimental conditions in healthy subjects. Two groups underwent a standardized online intervention, whereas the control group did not receive an intervention. While both intervention groups received the same online intervention, they were given different treatment rationales for the intervention: Participants in the causality group were told that the intervention will have beneficial effects on mood in the long-run, participants in the reversed causality group were instructed that mood will influence how they will perform in the intervention (see Appendix for details).

The intervention consisted of an online version of the expressive writing paradigm (Pennebaker & Beall, 1986), due to its feasibility and proven efficacy to reduce distress in healthy student and clinical populations (e.g. Baikié, Geerligs, & Wilhelm, 2012; Beyer et al., 2014). It needs to be noted that expressive writing leads to short-term deterioration of affect, followed by long-term increase in well-being (Sloan & Marx, 2004). Subjects in both intervention groups wrote about their most traumatic experience on three consecutive days for 20 minutes on each day.

Procedures

In total, participants were imbedded in the study for 46 days. The study encompassed three intervention days (days 1 to 3) and four assessments (days 1, 4, 10 and 46 or baseline, post intervention, mid-term and long-term follow-up, respectively). On intervention and assessment days participants received an email with an individual access code for a webpage differing in content according to group assignment and study day. For all groups, the website contained the outcome (PANAS, see below). On the intervention days, the website also contained an embedded video for the two intervention groups, showing a professional speaker explaining the rationale of the intervention. While the instruction was equal for both intervention groups in terms of structure and format, the treatment rationale, i.e. proposed meaning, differed in content between these groups (see Appendix for details). For the control group, this website contained the outcome only. On intervention days, participants of the intervention

groups completed the outcome before and after the expressive writing intervention, while the control group completed the outcome only once on each of the intervention days. The intervention, the different instructions as well as the assessments were conducted online in order to control for patient-therapist interaction. Participants were not informed about the existence of and the difference between the three groups. Subjects were debriefed after study participation. Participants were provided with written information about the study and written informed consent to participate in this study was obtained online and stored digitally. The protocol and consent procedure was approved by the institutional review board of the Department of Psychology of the University of Basel. The trial was not registered in a WHO approved registry before enrolment, because the trial initially was not considered to qualify as a clinical trial and comparable trials have not been registered as clinical trials (e.g. Langens and Schüler, 2005, 2007; Baikie et al., 2012; Beyer et al., 2014). However, the study was retrospectively registered at the German Clinical Trials Register (Identifier: DRKS00009428). The authors confirm that all ongoing and related trials for this intervention are registered. There were no changes to methods after trial commencement.

Measures

Possible effects of the experimental conditions were repeatedly assessed for up to 46 days after baseline with the Positive and Negative Affect Schedule (PANAS, Krohne, Egloff, Kohlmann, & Tausch, 1996), which was *a priori* defined as the primary outcome. The PANAS contains two scales with overall 10 five-point items assessing positive (e.g. "interested", "proud") and negative affect (e.g. "upset", "ashamed"). The internal consistency estimate of reliability for the PANAS scales in the total sample was good (positive affect: Cronbach's $\alpha=0.87$, negative affect: Cronbach's $\alpha=0.79$). In the intervention groups, linguistic content of written text, subjective rating of the severity of the reported traumatic experiences and the plausibility of their respective treatment rationale were employed to assess the validity of the intervention and experimental manipulation of treatment rationale. The linguistic content of the written text was analyzed with *Linguistic Inquiry and Word Count*, which quantifies words in a given text according to preset categories with a 70-80% hit rate (Wolf et al., 2008). For the purposes of this study, we used the word count of first person singular personal pronouns (e.g. *I, me, mine*), of negative and positive emotions (e.g. *sad, hate, worthless* and *happy, pretty*,

good), of cognitive processes (e.g. *distinguish, because, know*), of causality (e.g. *argument, influence, effect*) and of insight (e.g. *recognize, conscious, decision*). Subjective rating of the severity of the traumatic experiences was assessed with a single sentence 5-point item (i.e. *In general, how distressing is the experience you have just written about for you?*) and plausibility of the treatment rationale was operationalized with a 5-point item for each group (causality group: *Writing about a traumatic experience influences my well-being*, reversed causality group: *My well-being influences how I write about a traumatic experience*). All described measures were assessed online and stored in internal and secured data carriers.

Subjects

Subjects were recruited amongst psychology students at two Swiss Universities (University of Zurich and University of Basel) through mailing lists and web postings at both universities. Inclusion criteria were (1) ages of 18 years and older, (2) absence of any mental disorder by self-report, (3) not receiving psychological or psychiatric or medical treatment in the last six months by self-report and (4) a Toronto Alexithymia Scale score below 54 (Kupfer, Brosig, & Braehler, 2000), since alexithymia has been shown to influence effects of expressive writing (Paez, Velasco, & Gonzalez, 1999). Of 183 enrolled subjects, 19 did not fulfill inclusion criteria and were therefore excluded. A total of 164 subjects were randomly assigned to the three experimental groups and 112 participants completed the intervention and all assessments. Random allocation sequences were achieved with individual assignment codes in separate envelopes, opened by study personnel after confirmation of eligibility and informed consent. Eight subjects dropped out after randomization and before post-intervention assessment and 30 subjects were excluded due to low commitment in the expressive writing task, defined as writing for less than 15 minutes on intervention days by self-report, so that overall 126 subjects were considered per protocol (control group N=55, causality group N=36, reversed causality group N=35). Furthermore, 14 subjects did not complete all assessments after post-assessment, so that two samples were analyzed (completers N=112 and ITT N=126, see Figure 1). Subjects not meeting inclusion criteria as well as those excluded due to low commitment, i.e. not being considered as per protocol, did not differ significantly in any baseline demographic or baseline parameters (all $p > 0.40$). Mean age of participants was 23 years and 8 months (control group 22 years and 9 months; causality group 24 years

and 11 months; reversed causality group 23 years and 11 months) and the gender ratio was 86 women and 26 men (control group 37/11; causality group 25/7; reversed causality group 24/8). Overall, only 2 participants (1.8%) kept a regular diary. Participants received study credits for their participation. The flow of participants through the study is shown in Figure 1.

Data analysis

Based on assumed small to medium effects of our experimental manipulation of treatment rationale on intervention effects, a priori sample size calculation with the statistical software G*Power 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009) led to an optimal sample size of $N = 120$ ($f = 0.2$, 80% power, 5% alpha error, 3 groups, 4 assessments and correlation among repeated measures = 0.3). We assumed that our restrictive exclusion criteria, the population under investigation and the emotionally demanding task would lead to a substantial dropout of at least 30%. We therefore set out to recruit $N=180$. SPSS 21 statistical software (SPSS, Chicago, Illinois, USA) for Apple OS X was used for all statistical analyses. Analysis of variance or chi-square tests were used to examine demographic and clinical variables at baseline. A time by group by scale multivariate analysis of variance with subsequent time by group univariate analysis of variance for single PANAS scales was used to investigate differences between groups regarding treatment effects over time. For significant results between groups, Cohens f (calculated from SPSS partial eta-squared with the program G*Power 3.1: 0.1=small, 0.25=medium and 0.45=large) was used for time by group interaction effects. Within- and between-group effect sizes were calculated using Cohen's d (0.2=small, 0.5=medium, 0.8=large). All results are displayed as mean values and standard deviation unless otherwise indicated. All analyses were performed on per protocol subjects, i.e. with sufficient commitment in the writing task- All analyses of on protocol subjects were performed on an intent-to-treat (ITT) basis including randomized protocol subjects with baseline and at least post-intervention values (day 4) for the variable being analyzed (see Figure 1).

RESULTS

Characteristics of participants and validation check of intervention

Experimental groups did not differ significantly regarding gender distribution (female/male: causality group: 25/7, reversed causality group: 24/8 and control group: 37/11, chi-square=0.09, $p=0.96$, age

($F(2/109)=1.2$, $p=0.32$), baseline PANAS scores (Table 3, $F(4/218)=0.3$, $p=0.88$) and Toronto Alexithymia Scale total score ($F(2/109)=1.5$, $p=0.22$; control group: 38.0 (6.9), causality group: 38.6 (6.3) and reversed causality group: 35.9 (6.0)). Further, intervention groups did not differ substantially in linguistic content of written texts during the three intervention days, but participants in the reversed causality group used more first person singular personal pronouns in comparison to participants in the causality group on the first intervention day (Table 1). The causality and reversed causality groups reported comparable severity of the traumatic experiences on all three days and did not differ substantially in their ratings of the plausibility of treatment rationale (Table 2). The mean number of reported experienced traumatic events or troubling issues during the three intervention days was 1.9 (0.9) (one event: $N=28$, two events: $N=17$ and three events: $N=19$). The three most prevalent topics were: *Death or severe illness of family members* (e.g. cancer of father, 25%), *parental relationship problems* (e.g. divorce of parents, 22%) and *personal relationship problems* (e.g. adultery of spouse, 18%).

Insert tables 1 and 2 here

The expressive writing intervention led to significant decreases in positive as well as significant increases in negative affect (pre-post time effect: $F(2/61)=15.7$; $p<0.001$, $f=0.72$, Table 3), which both gradually attenuated over time (intervention day by pre-post interaction effect $F(4/248)=4.5$, $p=0.002$, $f=0.27$). Both intervention groups did not differ in their change of affect following the intervention (pre-post by group interaction effect: $F(2/61)=0.4$, $p=0.66$). Since participants in the control group were not subjected to the expressive writing intervention and thus a pre- and post-assessment on each day was not feasible, single assessments of the PANAS on days 1 to 3 were performed and compared with pre-intervention scores of the two intervention groups using a repeated multivariate analysis (Table 3). The PANAS scores in the control group did not change over time ($F(4/188)=0.9$, $p=0.45$).

Insert Table 3 here

Effects of the experimental manipulation

To test for possible effect of our experimental manipulation of the treatment rationale in the expressive writing intervention, short-, mid- and long-term effects were examined using repeated multivariate analyses between all three groups.

To assess short-term effects, differences in PANAS scores between day 1 (pre-assessment) and day 4 (1 day post-intervention) were calculated for the three groups. There was no change in PANAS scores over time (time effect: $F(2/108)=0.5$, $p=0.60$) or between groups (time by group interaction effect: $F(4/218)=0.2$, $p=.94$). To assess mid-term effects, group differences between day 1 to day 10 (7 days post-intervention) were calculated. Again, there were no changes in PANAS scores over time (time effect: $F(2/108)=0.2$, $p=0.80$) or between groups (time by group interaction effect: $F(4/218)=0.1$, $p=.98$). However, regarding long-term effects, differences in PANAS scores between day 1 to day 46 were significant (time effect: $F(2/108)=3.3$, $p=0.04$, $f=0.25$) and groups differed in the long-term changes of positive and negative affect (time by group interaction effect: $F(4/218)=2.9$, $p=0.02$, $f=0.23$). Significant group differences were found for both positive ($F(2/108)=3.6$, $p=0.03$, $f=0.25$) and negative affect ($F(2/108)=4.2$, $p=0.02$, $f=0.27$; see Figure 2). Post-hoc univariate comparisons on PANAS scores on day 46 indicated that with regard to positive affect, interventions groups did not differ significantly from each other, but both differed significantly from the control group (causality group vs. control group: $F(1/78)=4.2$, $p=0.04$, $f=0.23$; reversed causality group vs. control group: $F(1/78)=4.6$, $p=0.04$, $f=0.24$; causality group vs. reversed causality group: $F(1/62)=0.02$, $p=0.98$). With regard to negative affect the reversed causality group did not differ from the control group, but both groups differed from the causality group (causality group vs. control group: $F(1/78)=4.6$, $p=0.04$, $f=0.24$; reversed causality group vs. control group: $F(1/78)=0.03$, $p=0.88$; causality group vs. reversed causality group: $F(1/62)=4.0$, $p=0.05$, $f=0.25$). Standardized mean differences for significant effects were of small to medium size for positive affect (standardized between-groups differences at follow-up: causality group versus control group, $d=0.49$; reversed causality group versus control group, $d=0.50$; standardized within-group baseline to follow-up differences: causality group, $d=0.55$; reversed causality group, $d=0.37$; control group, $d=0.01$) and for negative affect (between groups at follow-up: causality group versus control group, $d=0.57$; causality group versus reversed causality group, $d=0.64$;

standardized within group baseline to follow-up differences: causality group, $d=0.44$; reversed causality group, $d=-0.22$; control group, $d=-0.19$).

Insert Figure 2 here

Analyses on ITT sample

Results of analyses on participants who were on treatment and completed at least the day 4 (post-intervention) assessment did not differ from the completers sample with regard to sample characteristics of participants, validation check of intervention as well as the magnitude of the short- and mid-term changes of PANAS scores between groups (data not shown). Also, the results for the long-term data of ITT sample was comparable to those of the completers sample, with minor changes of type one error for the time effect for PANAS scores and group by time interaction effect for positive affect.

Repeated multivariate analysis of group differences in PANAS scores from day 1 to day 46 indicated that PANAS scores changed over time (time effect: $F(2/122)=2.7$, $p=0.07$, $f=0.21$) and that groups differed in the long-term changes of positive and negative affect (time by group interaction effect: $F(4/246)=3.3$, $p=0.01$, $f=0.23$), with group differences with regard to positive ($F(2/122)=2.8$, $p=0.06$, $f=0.21$) and negative affect ($F(2/122)=4.7$, $p=0.01$, $f=0.28$). Post-hoc univariate comparisons of PANAS scores on day 46 indicated that with regard to positive affect, interventions groups did not differ from each other, but both differed from the control group (causality group vs. control group: $F(1/88)=4.2$, $p=0.04$, $f=0.21$; reversed causality group vs. control group: $F(1/82)=3.6$, $p=0.06$, $f=0.2$; causality group vs. reversed causality group: $F(1/68)=0.02$, $p=0.97$). With regard to negative affect the reversed causality group did not differ from the control group, but both groups differed from the causality group (causality group vs. control group: $F(1/88)=6.3$, $p=0.01$, $f=0.27$); reversed causality group vs. control group: $F(1/82)=0.5$, $p=0.51$; causality group vs. reversed causality group: $F(1/68)=9.6$, $p=0.003$, $f=0.37$). Standardized mean differences for significant effects in the ITT sample were of small to medium size for positive affect (standardized between-groups differences at follow-up: causality group versus control group $d=0.41$; reversed causality group versus control group $d=0.43$; standardized within-group baseline to follow-up differences: causality group $d=0.46$; reversed causality group $d=0.30$; control group $d=0.01$) and for negative affect (between groups at follow-up: causality

group versus control group, $d=0.57$; causality group versus reversed causality group, $d=0.64$; standardized within group baseline to follow-up differences: causality group, $d=0.33$; reversed causality group, $d=-0.28$; control group, $d=-0.2$).

DISCUSSION (WORD COUNT: 1174)

The aim of this randomized-controlled trial was to test the feasibility of an open/hidden paradigm in psychotherapy research and to assess the effect of the assumingly incidental treatment constituent of providing a treatment rationale to the assumingly characteristic psychotherapeutic intervention of expressive writing.

First, both the intervention as well as the experimental manipulation proved to be valid and feasible, as indicated by the linguistic content analysis, the expected detrimental effects of expressive writing on affect during intervention days and participants rating of plausibility of their treatment rationale. Noteworthy, intervention groups did not differ in these parameters, so that it is unlikely that observed effects of the experimental manipulation of the treatment rationale are the consequence of differences in the perception, implementation and direct effects of the intervention itself. Second, while the experimental manipulation did not influence affect in short- and mid-term, we observed significant long-term differences between groups. Both intervention groups reported higher positive affect six weeks after the intervention in comparison to the no-treatment control group. In contrast, while the causality group showed a decrease of negative affect, the reversed causality group did not seem to benefit from the intervention in the same fashion. All of these group differences were of medium effect size and the differences between the reversed causality and causality group in negative affect at follow-up exceeded the magnitude of observed within group-changes between baseline and follow-up.

In the following, the observed effects will be discussed from an expressive writing, a theoretical and a methodological perspective. With regard to the former, the results of the causality group followed the expected direction and time course, with a temporally delayed and medium-sized improvement in affect (e.g. Langens and Schüler, 2005, 2007; Beyer et al., 2014). Also, immediate deterioration of affect with long-term improvement is proposed as the normal psychokinetic of expressive writing interventions (Baikie et al., 2012) and follow-up duration has been found to be a significant predictor

of expressive writing effects, with longer-term follow-up studies showing better outcome than studies with shorter-term follow-up (Travagin, Margola, & Revenson, 2015). It needs to be noted that although the observed effects are comparable to previous publications (see above), our trial is at best considered an efficacy trial and thus our results do not shed light on the effectiveness of online expressive writing interventions.

From a theoretical perspective, the observed response differences between the intervention groups are per se not unexpected. For example, a recent within-subject randomized blinded balanced placebo design observed higher analgesic effects with overt in contrast to covert lidocaine injections in healthy subjects (Lund et al., 2014). Similarly, providing an otherwise similar aerobic exercise training or a work-related behavior with group-specific therapeutic rationales led to larger changes in self-esteem (Desharnais et al., 1993) or physiological health parameters (Crum & Langer, 2007), respectively. However, it needs to be noted that the reversed causality group benefitted from the intervention with regard to positive, but not to negative affect. Here, an early study on the effects of either open or hidden administration of a pharmacological intervention is informative. Ross et al. (1962) observed discomfort in subjects receiving 10 mg of d-amphetamine covertly, while the overt administration resulted in neutral to comfortable mood, arguing that subjects who covertly received the pharmacological intervention, "had no 'therapeutic' set and presumably reported their feelings 'honestly' on the mood scales" (Ross et al., 1962, p. 391). However, the processes underlying the observed differences, such as differences in experiencing or attributing symptoms or both, need to be elucidated in further studies. Regarding the methodological approach, the observed effects of our trial support the feasibility and utility of an open/hidden administration to disentangle effects of incidental treatment constituents from those of characteristic treatment constituents. In this regard, the beneficial effects on positive affect in both intervention groups could be seen as being caused by characteristic constituents of expressive writing, while the incidental treatment constituent of providing a treatment rationale is needed to obtain full effects, i.e. beneficial changes in both positive and negative affect. It needs to be noted that the definition of what is to be considered characteristic or incidental in psychotherapy is complicated by the fact that many treatments include the provision of a treatment rationale without defining its role in the treatment theory as characteristic. While this in general highlights the need to revise treatment

theories (Gaab et al., 2015), the definition of what is to be considered incidental or characteristic in the case of expressive writing was considerably facilitated by the fact that the original instruction (Pennebaker, 2000) did not contain the provision of a treatment rationale and thus is clearly considered incidental to the treatment. According to Grünbaum's conceptualization (1981), a treatment qualifies as *verum*² when at least one characteristic treatment constituent is therapeutic for a given disorder. With regard to our findings, we have reason to assume that expressive writing qualifies as *verum*, since its effects are observable even when administered covertly. However, there is an important amendment to this conclusion. Explanatory models and theories of expressive writing currently focus on emotion-related processes, i.e. propose disinhibition, emotional catharsis, cognitive-processing, self-regulation, exposure, social integration or development of a coherent narrative as underlying principles of change (Baikie et al., 2012; Frattaroli, 2006). Based on our results and given that 1) expressive writing effects are strongly associated with positive response expectancies, 2) a full confrontation with the traumatic experience is not a prerequisite for these effects (Langens & Schuler, 2007) and 3) that writing about traumatic events is as beneficial as writing about positive events (Frattaroli, 2006), future research is warranted to elucidate characteristic treatment constituents in expressive writing and to revise the underlying treatment theory accordingly.

There are several aspects of our study that need to be considered critically. First, participants in the reverse causality group may have known about the beneficial effects of writing. However, controlling for this is difficult, if not impossible, as asking participants about their assumptions regarding possibly positive effects of writing might itself provoke or prime such assumptions. Therefore, we opted against a direct inquiry and instead assessed the plausibility of the respective condition rationale. Our results showed no difference between the intervention groups, which indicates that participants found both possible causal directions between affect-change and writing equally acceptable. Second, we experimentally manipulated the treatment rationale and thus provided the causality group, i.e. the open administration of the intervention, with a description of expected positive effects of the intervention. In the original instructions (e.g. Pennebaker, 2000) this clear description and direction of effects is missing. However, the therapeutic properties of expressive writing interventions are inherently tied to

² Which he chooses to label "non-placebo"

accounts of this intervention (e.g. Pennebaker, 2011) and even with the original instructions, i.e. without providing a description of expected positive effects, high positive expectancies that the writing intervention will have beneficial effects were found in comparable populations (Langens & Schuler, 2007). Also – at least with respect to ethical guidelines in psychotherapy – providing a treatment rationale is integral to normative guidelines (see Blease et al., 2016). Since our aim was to test the effects of an experimental manipulation of providing a treatment rationale (and not to replicate the effects of the original expressive writing instructions), we choose to make the implicit explicit, by providing and manipulating the rationale of the two conditions. However, it needs to be acknowledged that our instructions differed from those in Pennebaker (2000), which did not provide a description of any anticipatory effect or treatment rationale. Third and closely related to the previous aspect, our hidden administration of the online expressive writing intervention differs from similar approaches in medical research, where the patient is unaware when a given substance is administered. Instead, we operationalized the hidden administration by providing an alternative non-therapeutic rationale to ensure that both experimental conditions do not differ in their credibility and to minimize the possibility that the expressive writing intervention would per se be perceived as therapeutic. Given that this has been observed in trials not providing a rationale (e.g. Langens & Schüler, 2007), we addressed this problem by implementing a non-therapeutic rationale. Finally, it needs to be noted that we carried out a large number of analyses and thus there is a probability that our main finding is due to chance. However, it needs to be considered that 1. we carefully analyzed pre-defined hypotheses with multivariate methods, 2. our finding of a significant short-term deterioration and a long-term improvement replicates previous findings on expressive writing from independent groups and 3. our main finding of a significant group difference in the course of long-term negative affect is of medium effect size and the consequence of a well-controlled experimental design. Therefore, we consider the probability of a chance finding as minimal

To conclude, we consider the manipulation of assumingly incidental treatment constituents while keeping assumingly characteristic treatment constituents constant, as a promising design for psychotherapy research. Although our study might also illustrate the difficulties of this approach, it is far from new. Faced with the problem of an effective treatment without a proven mechanism, Benjamin

Franklin and colleagues employed probably the first documented open/hidden trial to test Anton Mesmers animal magnetism at the end of the 18th. century. Using a specially designed blindfold, the servant of an experienced practitioner of magnetism was put to test: "Magnetized next with eyes uncovered, he feels tingling in his forehead when the metal rod is brought close to it; blindfolded again, he feels no tingling when the rod is brought close" (cited from Franklin et al., 2002; p. 347). These results, amongst others, convinced the Franklin commission that "having finally demonstrated by decisive experiments that the imagination without magnetism produces convulsions, & that magnetism without imagination produces nothing" or state otherwise, that the otherwise impressive effects of magnetism was fully driven by its incidental and not by any characteristic treatment constituents.

REFERENCES

- Ahn, H. N., & Wampold, B. E. (2001). Where oh where are the specific ingredients? A meta-analysis of component studies in counseling and psychotherapy. *Journal of Counseling Psychology, 48*(3), 251-257.
- American Psychological Association, Division 12. Research-Supported Psychological Treatments. <http://www.psychologicaltreatments.org>, retrieved 28. July, 2015
- Baikie, K. A., Geerligs, L., & Wilhelm, K. (2012). Expressive writing and positive writing for participants with mood disorders: an online randomized controlled trial. *J Affect Disord, 136*(3), 310-319.
- Baskin, T. W., Tierney, S. C., Minami, T., & Wampold, B. E. (2003). Establishing specificity in psychotherapy: A meta-analysis of structural equivalence of placebo controls. *Journal of Consulting and Clinical Psychology, 71*(6), 973-979.
- Beecher, H. K. (1955). The powerful placebo. *J Am Med Assoc, 159*(17), 1602-1606.
- Bell, E. C., Marcus, D. K., & Goodlad, J. K. (2013). Are the parts as good as the whole? A meta-analysis of component treatment studies. *J Consult Clin Psychol, 81*(4), 722-736.
- Beyer, J. A., Lumley, M. A., Latsch, D. V., Oberleitner, L. M., Carty, J. N., & Radcliffe, A. M. (2014). Computer-based written emotional disclosure: the effects of advance or real-time guidance and moderation by Big 5 personality traits. *Anxiety Stress Coping, 27*(5), 477-493.

- Blease, C., Trachsel, M., grosse Holtforth, M. (2016). Paternalism, Placebos, and Informed Consent in Psychotherapy: The Challenge of Ethical Disclosure. *Verhaltenstherapie* 2016, epub January 28, 2016, DOI: 10.1159/000442928
- Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *J Consult Clin Psychol*, 66(1), 7-18.
- Crum, A. J., & Langer, E. J. (2007). Mind-set matters: exercise and the placebo effect. *Psychol Sci*, 18(2), 165-171.
- Cuijpers, P., Driessen, E., Hollon, S. D., van Oppen, P., Barth, J., & Andersson, G. (2012). The efficacy of non-directive supportive therapy for adult depression: a meta-analysis. *Clin Psychol Rev*, 32(4), 280-291.
- Desharnais, R., Jobin, J., Cote, C., Levesque, L., & Godin, G. (1993). Aerobic exercise and the placebo effect: a controlled study. *Psychosom Med*, 55(2), 149-154.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav Res Methods*, 41(4), 1149-1160.
- Fergusson, D., Glass, K. C., Waring, D., & Shapiro, S. (2004). Turning a blind eye: the success of blinding reported in a random sample of randomised, placebo controlled trials. *BMJ*, 328(7437), 432.
- Fluckiger, C., Del Re, A. C., & Wampold, B. E. (2015). The sleeper effect: Artifact or phenomenon-A brief comment on Bell et al. (2013). *J Consult Clin Psychol*, 83(2), 438-442; discussion 443-434.
- Frank, J. D. (1986). Psychotherapy--the transformation of meanings: discussion paper. *J R Soc Med*, 79(6), 341-346.
- Franklin, B., Majault, Le, R., Sallin, Bailly, J. S., D'Arcet, et al. (2002). Report of the commissioners charged by the King with the examination of animal magnetism. 1784. *Int J Clin Exp Hypn*, 50(4), 332-363.
- Fratraroli, J. (2006). Experimental disclosure and its moderators: a meta-analysis. *Psychol Bull*, 132(6), 823-865.

- Gaab, J., Blease, C., Locher, C., & Gerger, H. (2015). Go Open: A Plea for Transparency in Psychotherapy. *Psychology of Consciousness: Theory, Research, and Practice*. Advance online publication. <http://dx.doi.org/10.1037/cns0000063>
- Grünbaum, A. (1981). The placebo concept. *Behav Res Ther*, *19*(2), 157-167.
- Jones, D. S., & Podolsky, S. H. (2015). The history and fate of the gold standard. *Lancet*, *385*(9977), 1502-1503.
- Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annu Rev Clin Psychol*, *3*, 1-27.
- Kazdin, A. E., & Wilcoxon, L. A. (1976). Systematic desensitization and nonspecific treatment effects: a methodological evaluation. *Psychol Bull*, *83*(5), 729-758.
- Krohne, H. W., Egloff, B., Kohlmann, C. W., & Tausch, A. (1996). Untersuchungen mit einer deutschen Version der "Positive and Negative Affect Schedule (PANAS)". *Diagnostica*, *42*, 139-156.
- Kupfer, J., Brosig, B., & Brahler, E. (2000). [Testing and validation of the 26-Item Toronto Alexithymia Scale in a representative population sample]. *Z Psychosom Med Psychother*, *46*(4), 368-384.
- Langens, T. A., & Schuler, J. (2005). Written emotional expression and emotional well-being: the moderating role of fear of rejection. *Pers Soc Psychol Bull*, *31*(6), 818-830.
- Langens, T. A., & Schuler, J. (2007). Effects of written emotional expression: the role of positive expectancies. *Health Psychol*, *26*(2), 174-182.
- Locher, C., Hasler, S., Gaab, J. (2016). When Do Placebos in Psychotherapeutic Research Work? A Systematic Review on the Example of Systematic Desensitization. *Verhaltenstherapie, epub January 28, 2016, DOI: 10.1159/000443464*
- Lund, K., Vase, L., Petersen, G. L., Jensen, T. S., & Finnerup, N. B. (2014). Randomised controlled trials may underestimate drug effects: balanced placebo trial design. *PLoS One*, *9*(1), e84104.
- Marcus, D. K., O'Connell, D., Norris, A. L., & Sawaqdeh, A. (2014). Is the Dodo bird endangered in the 21st century? A meta-analysis of treatment comparison studies. *Clin Psychol Rev*, *34*(7), 519-530.

- Moerman, D. E., & Jonas, W. B. (2002). Deconstructing the placebo effect and finding the meaning response. *Ann Intern Med*, 136(6), 471-476.
- Munder, T., Fluckiger, C., Gerger, H., Wampold, B. E., & Barth, J. (2012). Is the Allegiance Effect an Epiphenomenon of True Efficacy Differences Between Treatments? A Meta-Analysis. *Journal of Counseling Psychology*, 59(4), 631-637.
- Paez, D., Velasco, C., & Gonzalez, J. L. (1999). Expressive writing and the role of alexythimia as a dispositional deficit in self-disclosure and psychological health. *J Pers Soc Psychol*, 77(3), 630-641.
- Pennebaker, J. W. (2000). Writing and health: Some practical advice. Retrieved 28. July, 2015
- Pennebaker, J. W. (2011, 24.8.2011). The expressive writing method. Retrieved 28. July, 2015
- Pennebaker, J. W., & Beall, S. K. (1986). Confronting a traumatic event: toward an understanding of inhibition and disease. *J Abnorm Psychol*, 95(3), 274-281.
- Rosenthal, D., & Frank, J. D. (1956). Psychotherapy and the placebo effect. *Psychol Bull*, 53(4), 294-302.
- Ross, S., Krugman, A. D., Lysterly, S. B., & Clyde, D. J. (1962). Drugs and placebos: A model design. *Psychological reports*, 10, 383-392.
- Sloan, D. M., & Marx, B. P. (2004). A closer examination of the structured written disclosure procedure. *J Consult Clin Psychol*, 72(2), 165-175.
- Travagin, G., Margola, D., & Revenson, T. A. (2015). How effective are expressive writing interventions for adolescents? A meta-analytic review. *Clin Psychol Rev*, 36, 42-55.
- Wampold, B., & Imel, Z. E. (2015). *The great psychotherapy debate. Second edition*: Routledge.
- Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. N. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, "all must have prizes". *Psychological Bulletin*, 122(3), 203-215.
- Wolf, M., Horn, A., Mehl, M., Haug, S., Pennebaker, J. W., & Kordy, H. (2008). Computergestützte quantitative Textanalyse: Äquivalenz und Robustheit der deutschen Version des Linguistic Inquiry and Word Count [Computer-aided quantitative text analysis: Equivalence and

reliability of the Ger-man adaptation of the Linguistic Inquiry and Word Count]. *Diagnostica*,
1, 85-98.