

This paper has the same format as the 2022 paper - but with an important difference: in 2021 students had a 24-hour period to complete the paper, but in 2022 there is a time limit of 3 hours

UNIVERSITY COLLEGE LONDON

EXAMINATION FOR INTERNAL STUDENTS

MODULE CODE : **PSYC0185**

ASSESSMENT : **PSYC0185B8PC**
PATTERN

MODULE NAME : **PSYC0185 - Research Methods in Clinical Psychology**

LEVEL: : **Postgraduate**

DATE : **27/05/2021**

TIME : **10:00**

This paper is suitable for candidates who attended classes for this module in the following academic year(s):

Year
2020/21

Additional material	Peer-reviewed paper – Keinonen et al. (2019) Guide for Critiquing a Paper
Special instructions	
Exam paper word count	3,000

TURN OVER

DOCTORATE IN CLINICAL PSYCHOLOGY

Paper 2: Research Methods in Clinical Psychology

Instructions

- You will have a 24-hour time frame to complete the exam and upload it onto AssessmentUCL, as instructed. The exam deadline is **10:00 28th May 2021**.
- Please read the uploaded paper by Keinonen et al. (2019) and **answer all eight of the questions** below.
- We recommend that you read the paper carefully from start to finish before starting to answer the questions.
- The uploaded Guide for Critiquing a Paper form may be a useful tool to help you structure your thinking for some of the questions.
- If there is any important information missing from the paper that you think is relevant to answering one of the questions, you can reflect on this and its implications for the paper.
- You should **write no more than 3000 words** in total across the eight questions. Each **question should therefore be allocated up to 375 words**. All written answers beyond the 3000-word limit will not be marked.

1. Comment on the overall design of the study (including any conditions or time points that were chosen) and whether it is appropriate for the research hypotheses.

2. Comment on the sample and sampling procedure of the study. Where relevant consider inclusion and exclusion criteria, allocation procedures / randomisation and any potential sources of bias.

3. Comment on the dependent and independent variables. Where relevant consider levels of measurement, their reliability and validity, and their appropriateness given the research question.

4. What were the main findings reported? Comment on the specific analyses and results that supported these findings.

Continues on next page

5. To what extent were the conclusions valid? Where relevant, consider concepts such as statistical conclusion validity, effect sizes, and internal, external and ecological validity.
6. Discuss the ethical issues you would be considering if you were reviewing the introduction and methods sections as part of an application for ethical approval.
7. If you were to design a study to build on the findings of this paper, what might be a good topic / area for study that would be well suited to a qualitative approach? Justify your reasoning, i.e., why would a qualitative approach be particularly appropriate?
8. What qualitative approach and associated epistemology would you use and why?

END OF PAPER

Research Methods - Critiquing a Paper

Marc Tibber and Liam Mason, 2019

Read the abstract first and hold in mind the key conclusions drawn by the authors. Focus on evaluating these rather than evaluating the paper exhaustively.

1.
 - a) Does the report specify a hypothesis (explicit or implicit)? If so, what is it?
 - b) What background or rationale is provided as justification for any hypothesis?
2.
 - a) Categorise the dependent and independent variables.
 - b) Categorise type of measurement (nominal, interval... etc.). This will help to evaluate the analyses used.
 - c) What are the reliability and validity of the measures?
 - d) How well do the selected variables address the research hypothesis/hypotheses?
3.
 - a) What sampling methods were used?
 - b) Was there any risk of selection bias?
 - c) Are there any other potential sources of bias?
4.
 - a) What "design" was used? (Descriptive, correlational, quasi-experimental... randomised, etc.).
 - b) Do the method and design address the research hypothesis/hypotheses?
5.
 - a) What analyses / statistics were used? (Were they appropriate?).
 - b) Was there likely to be sufficient power?
6. What was the main result?
7. What conclusions were drawn?
8. Validity analysis:
 - a) Statistical conclusion validity (asks "Is there an effect?").
 - b) Internal validity (asks "Is it causal?").
 - c) Construct validity (asks "What does it mean?").
 - d) External validity (asks "Does it generalise?").
9. What are your main criticisms of the report?
10. What are the main strengths of the report?
11. What are the paper's broader implications, e.g. impact, ethical / social / political?



Long-term stability of early sudden gains in an acceptance and values-based intervention

Katariina Keinonen*, Heidi Kyllönen, Piia Astikainen, Raimo Lappalainen

Department of Psychology, University of Jyväskylä, P.O. Box 35, FI-40014, Finland

ARTICLE INFO

Keywords:

Early sudden gains
Long-term outcome
Depression
Novice therapists
Brief intervention

ABSTRACT

Though previous research has extensively reported that sudden gains are associated with superior treatment results, research on the long-term effects and stability of sudden gains is not as consistent. The current study explored the long-term stability of *early* sudden gains (ESGs) observed in a brief acceptance and values-based intervention for depression provided by novice therapists. The participants were 56 volunteers diagnosed with major depressive disorder. Among the participants, 23% experienced ESGs, i.e. they reached the status of improved or recovered in the Reliable Change Index (RCI; Jacobson & Truax, 1991) classification after only two sessions. The current study examined the level of depressive symptoms (BDI-II), psychological flexibility (AAQ-II), believability of depressive thoughts (ATQ-B), and hopefulness (ASHS) 6 and 12 months after the intervention. Of the original participants, 77% ($n = 43$) took part in the 12-month follow-up. The results showed that positive changes achieved during the two first sessions in the ESG group were maintained up to 12 months after the intervention. The ESG group remained improved or recovered in the RCI classification through the 12-month follow-up period. However, at the 12-month follow-up, there were only small differences between the ESG and the non-ESG groups. The results suggest that participants achieving ESGs show stable improvements lasting up to 12 months after the treatment.

1. Introduction

In the treatment of depression and various anxiety disorders, sudden gains—i.e. large, stable, and sudden changes in symptom severity—have been associated with superior treatment outcomes at post-treatment (e.g., Hedman et al., 2014; Keller, Feeny, & Zoellner, 2014; Norton, Klenck, & Barrera, 2010). A meta-analysis of 16 studies on sudden gains showed the effect size that sudden gains have on treatment outcomes is particularly large (Hedges' $g = 0.75$) when cognitive behavioral treatments are employed (Aderka, Nickerson, Bøe, & Hofmann, 2012). As the typical magnitude of sudden gains is relatively large—around 10–13 points on the Beck Depression Inventory (BDI; Beck & Steer, 1987; e.g., Tang, DeRubeis, Hollon, Amsterdam, & Shelton, 2007)—sudden gains appear to have high clinical relevance during the treatment.

Overall, sudden gains occurring at various stages of treatment are an interesting subtype of improvement not only due to the superior treatment results that they are associated with but also due to the high frequency at which they occur. Studies have typically reported that approximately 30–40% of participants experience at least one sudden

gain during the treatment (e.g., Hunnicutt-Ferguson, Hoxha, & Gollan, 2012; Tang et al., 2007). These results suggest that understanding sudden gains may help clinicians to identify and take advantage of this dramatic and frequent phenomenon.

One of the features of sudden gains that still needs to be clarified is how far reaching the effects of sudden gains are on treatment results. Whereas several studies have reported the effect of sudden gains immediately after treatment, reports of long-term treatment results are rarer. Some encouraging findings can be found, however. In the treatment of depression, sudden gains have been associated with lower rates of relapse and symptom recurrence during a two-year follow-up phase (Tang et al., 2007). In the treatment of anxiety disorders, it has been reported that sudden gains are associated with superior treatment results up to 12-month follow-up (Bohn, Aderka, Schreiber, Stangier, & Hofmann, 2013; Hedman et al., 2014). Similarly, Aderka, Appelbaum-Namdar, Shafran, and Gilboa-Schechtman (2011) found that children who had sudden gains while being treated for posttraumatic stress disorder had superior treatment outcomes at posttreatment, and they maintained the advantage at 12-month follow-up. In line with these results, a meta-analysis showed that sudden gains have a moderate

* Corresponding author.

E-mail addresses: katariina.keinonen@jyu.fi (K. Keinonen), heidimaria.kyllonen@gmail.com (H. Kyllönen), piia.s.astikainen@jyu.fi (P. Astikainen), raimo.lappalainen@jyu.fi (R. Lappalainen).

<https://doi.org/10.1016/j.jcbs.2019.06.006>

Received 18 December 2018; Received in revised form 3 June 2019; Accepted 25 June 2019

2212-1447/ © 2019 Association for Contextual Behavioral Science. Published by Elsevier Inc. All rights reserved.

effect at follow-up compared to those not reporting sudden gains during the treatment ($n = 8$ studies, Hedges' $g = 0.56$, pooled follow-up durations, mean $[M] = 4.44$ months; Aderka et al., 2012). Also, a study by Lemmens, DeRubeis, Arntz, Peeters, and Huibers (2016) suggested that sudden gains have prognostic value up to five months after treatment but not after 12 months. More studies are needed to draw conclusions on how stable the changes associated with sudden gains are and how far reaching the advantage is when compared to those not experiencing sudden gains.

Sudden gains can occur at any point during the treatment, but some studies have reported that their frequency is higher during the first few sessions (e.g., Hunnicutt-Ferguson et al., 2012; Masterson et al., 2014). For example, Dour, Chorpita, Lee, and Weisz (2013) observed that sudden gains were more likely to occur between the second and third sessions than at any other time, and most of the observed sudden gains occurred before session six. Similarly, Hunnicutt-Ferguson et al. (2012) found that the frequency of sudden gains was highest for the first session, and, again, most sudden gains occurred before session six.

The available studies that focus on *early* sudden gains (ESG) have reported somewhat mixed results on the effect on the treatment outcome. ESG is a similar concept to that of *rapid response*, that refers to clinically significant changes after the first four weeks of the treatment. However, ESG refers to large and clinically significant changes during these first four weeks of treatment. Whereas some have found that ESGs predict overall treatment outcome (e.g., Hunnicutt-Ferguson et al., 2012; Kelly, Roberts, & Ciesla, 2005), there are also studies that have found that ESGs do not predict better outcomes (Dour et al., 2013). In addition, Clerkin, Teachman, and Smith-Janik (2008) reported that sudden gains occurring after session one do not predict greater improvements, but sudden gains occurring after session two do predict greater improvements. In accordance with this, we have previously reported that sudden gains occurring before the third session predict a better overall treatment response in the treatment of depression as measured at post treatment (Keinonen, Kyllönen, Astikainen, & Lappalainen, 2018; Kyllönen et al., 2018). Finally, it has been suggested that ESGs (i.e. sudden gains occurring during the first third of the treatment) may be more significant for treatment results than sudden gains occurring later in the treatment (Kelly et al., 2005). Based on the available studies, further exploration of the clinical and theoretical importance of ESGs is needed.

In addition to the need for more studies with longer follow-ups, we also need more studies that investigate sudden gains in treatments provided by inexperienced therapists. If sudden gains consistently occur in treatments provided by novice therapists, this may suggest sudden gains are not related to extensive training or high competence. Indeed, Greenfield, Gunthert, and Haaga (2011) have reported that novice therapists working in a psychotherapy training clinic produced sudden gains at a comparable frequency (23% of participants experienced sudden gains) to that reported for experienced therapists. To our knowledge, there are no studies reporting the long-term stability of sudden gains in the treatments administered by novice therapists.

The current study aims to extend our understanding of sudden gains by exploring the effects of ESGs on long-term treatment results in an acceptance and values-based treatment (ACT) provided by novice therapists. We have previously reported that ESG was associated with superior treatment results at posttreatment in a six-session intervention delivered by student therapists without any prior clinical experience (Keinonen et al., 2018). All participants were diagnosed as having a major depressive disorder, and ESG was defined as being classified as improved or recovered according to the RCI classification (Jacobson & Truax, 1991) using Beck Depression Inventory-II scores after only two sessions (BDI-II; Beck, Steer, & Brown, 2004). We found that 23% of participants reached the status of early sudden gainers using the RCI method to identify ESG (Keinonen et al., 2018). Superior treatment outcomes were detected on both the level of depressive symptomatology and on the level of psychological flexibility, which was used as a

secondary measure for therapeutic change (Keinonen et al., 2018).

In the current study, we investigated whether significant changes in depression symptoms occurring at early stages of treatment (i.e. during the first two weeks) predict the long-term treatment outcome when participants diagnosed with a major depressive disorder were offered a six-session ACT intervention delivered by inexperienced therapists. We wanted to explore: (1) Are ESGs that are observed in depression symptoms within the two first sessions maintained after the intervention? Do those who show ESGs in depression show superior long-term treatment outcomes (at 6 and 12 months) on the levels of depressive symptomatology, psychological flexibility, hopefulness, and believability of depressive thoughts? (2) How strong is the association between ESGs and long-term treatment results (at 6 and 12 months)? (3) Are ESGs associated with the RCI classification status (recovered, improved, unchanged, or declined), the use of medication for depression, or psychological treatment for depression at the 6- and/or 12-month follow-up phases?

2. Method

2.1. Participants

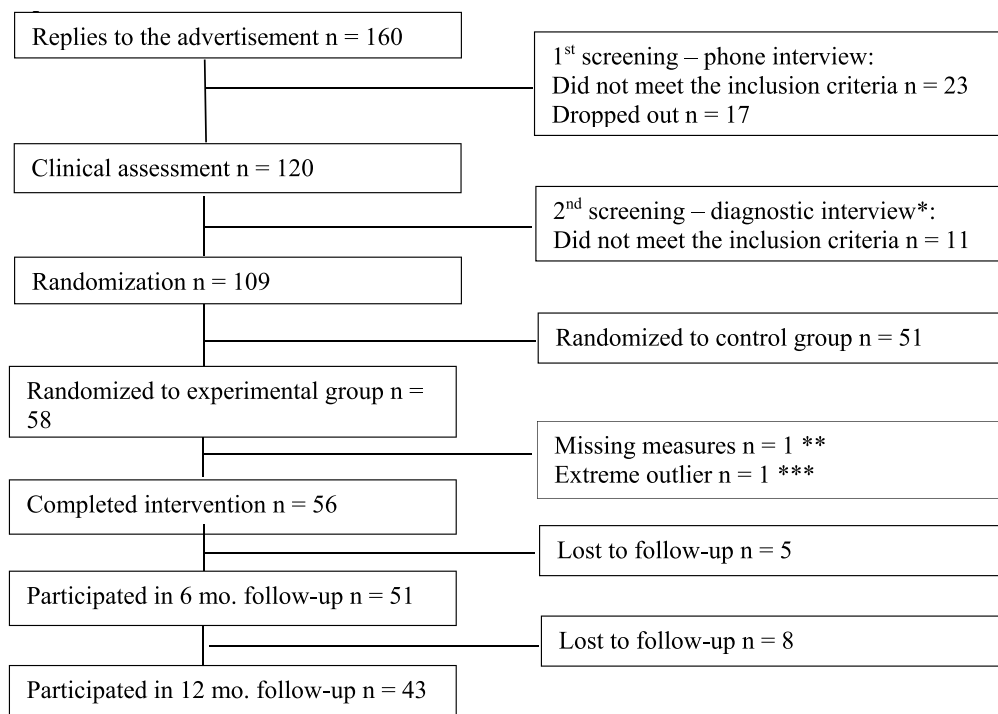
The current study is part of a larger study examining the efficacy of a six-session ACT intervention for major depressive disorder delivered by novice therapists (Kyllönen et al., 2018). The participants were volunteers with a diagnosis of major depressive disorder that was based on the criteria in the *International Classification of Diseases* (World Health Organization, 1992) and was obtained in an interview with a medical doctor as a part of the recruitment for the study. Exclusion criteria for the study were: (1) psychiatric diagnosis other than depression, (2) neurological diagnosis, (3) misuse of alcohol or drugs, and (4) on-going psychological treatment.

A total of 56 participants were included in the current study with 43 (77%) of the participants completing all measures up to the 12-month follow-up. The flow of participants is shown in Fig. 1. Forty-three (77%) of the participants were women, half (50%) were employed, and more than half (54%) were married or cohabiting. The mean age of the participants was 49.2 years (standard deviation $[SD] = 11.74$, range = 19–65). The majority of participants were diagnosed with mild depression (61%; single episode, mild $n = 18$; recurrent, mild $n = 16$). Dysthymic disorder was diagnosed for 7% of the participants ($n = 4$), and moderate depression was diagnosed for 32% (single episode $n = 9$; recurrent $n = 9$). For more detailed information, see Table 1. Table 1 also describes the participants classification based on the BDI-II.

The follow-up phases were 6 and 12 months after the six-session ACT intervention and included an interview and the completion of questionnaires. A 15–20 min interview was conducted face-to-face at the 6-month follow-up and by the phone at the 12-month follow-up. The participants were contacted by phone to invite them to participate in the follow-up phases.

2.2. Intervention

The therapists were master's students of psychology ($n = 37$) with no prior clinical experience (mean age 24 years, $SD = 3.8$, range 20–39, 92% female). Training consisted of four full-day lectures (32 h) on constructing a case formulation model and learning the principles and methods of ACT. The six-session intervention was a semi-structured ACT-based intervention delivered by student therapists with no prior clinical experience. The first two sessions were highly structured. During the first session, the therapist interviewed the client to get an overview of the client's situation and constructed an individual problem list. At the end of the session, the therapist explained and presented a home assignment regarding values. The second session included the presentation of a case formulation model (Haynes & O'Brien, 2000) based on the interview and a discussion of the model. After discussing



Note: * = Screening interview carried out by a physician. ** = Baseline BDI-II missing. *** = BDI-II change across two sessions from 40 to 2 points.

Fig. 1. Flow of participants. Note: * = Screening interview carried out by a physician. ** = Baseline BDI-II missing. *** = BDI-II change across two sessions from 40 to 2 points.

Table 1
Descriptive data for background and baseline variables.

	N	%
Sex	13	23
Male	43	77
Female		
Age group	5	9
18–29	18	32
30–49	33	59
50–65		
Marital status	10	18
Single	30	54
Married/Cohabiting	15	27
Divorced/Widowed	1	2
N/A		
Education	2	4
Basic education	38	68
Upper secondary education	16	29
Higher education		
Antidepressive medication	20	36
Yes	36	64
No		
Previous psychotherapy	8	14
Yes	48	86
No		
BDI-II classification	19	34
Mild	20	36
Moderate	17	30
Severe		

the model, the therapist reviewed the homework assignment regarding values, and the case formulation model was re-evaluated based on the value discussion. The client was presented with an individually chosen metaphor, the observer exercise (Hayes, Strosahl, & Wilson, 1999), and

a mindfulness exercise (i.e., “Follow your breath”; Hayes et al., 1999). As a homework assignment, the client was instructed to take value-based actions and practice mindfulness for 5–10 min per day. The last four sessions could be more individualized, but the student therapists were instructed to use exercises and metaphors in each session with 12 exercises being obligatory, though the timing was freely chosen. The student therapists received weekly group supervision (2 h, total of 10 h) with an expert in ACT. The supervisor helped the students to plan the intervention based on the case conceptualization. In addition to the group supervision, the student therapists participated in weekly peer-supervision groups. A more detailed description of the intervention can be seen in Keinonen et al. (2018) and Kyllönen et al. (2018).

2.3. Treatment integrity

The student therapists' adherence and competence were evaluated using the ACT adherence scale (Plumb & Vilardaga, 2010). The ACT Adherence scale evaluates competence and adherence on a scale of 1–5, where a rating of 1 for competence suggests “The therapist did not competently address any of the client's needs, did not attend to the client's responses to treatment targets, and did not apply any of the processes outlined in the manual.” A rating of 5 for competence suggests “The therapist consistently addressed the client's needs, consistently attended to the client's response to treatment targets, and applied the processes outlined in the manual very clearly and in-depth”. For adherence, a rating of 1 suggests “The session was entirely off topic or focused entirely on general assessment without addressing any of the other processes outlined in the therapist manual.” A rating of 5 for adherence suggests “The therapist spent most of the session doing a general assessment of functioning and applied more than one of the therapy processes in an extremely in-depth manner”. The process-

specific subscales are also rated on a scale of 1–5 based on frequency and depth, where a rating of 1 indicates that “The variable never explicitly occurred.” and a rating of 5 indicates that “The variable occurred with great frequency and was addressed by the therapist in a very in-depth manner.”

To evaluate treatment integrity, two randomly selected sessions were coded from each therapist ($n = 37$) providing a total of 74 coded sessions. A psychologist with clinical experience of using ACT methods coded the material. Three other coders coded a sample to evaluate interrater reliability; two very experienced ACT-psychotherapists were assigned to code 7 sessions each, and a psychology student was assigned to code 22 sessions. The intraclass correlation coefficient (ICC) for absolute agreement between each coder pair reflected good interrater reliability: the average ICC was 0.80–0.84 for the three coder pairs over all variables. The student therapists delivered the intervention with satisfying levels of adherence and competence (on a scale of 1–5, $M = 3.31$, $SD = 0.87$ and $M = 3.23$, $SD = 0.74$, respectively), though not on a professional level. On the ACT Adherence Scale, a rating of 3 for competence suggests “The therapist sometimes addressed the client’s needs, sometimes attended to the client’s response to treatment targets, and applied the processes outlined in the manual only superficially” and a rating of 3 for adherence suggests “The therapist spent at least half of the session attending to at least one of the processes outlined at any point in the therapy manual, also attending to general assessment, in a somewhat in-depth manner.” The descriptive statistics for all the ACT process variables that were included in coding are reported in Table 2.

2.4. Measures

Four measures were used in the current analysis to evaluate the level of symptoms and psychological processes at four timepoints: (1) at pretreatment, (2) after two sessions (2–4 weeks after the pre-measurement), (3) at six months after the intervention, and (4) at 12 months after the intervention.

Beck Depression Inventory II (BDI-II). Changes in depressive symptomatology were evaluated with the BDI-II (Beck et al., 2004). The BDI-II is a 21-item questionnaire. Each item asks the client to choose one of four statements that correspond to increasing severity of symptoms of depression (e.g., 0 = I do not feel sad, 1 = I feel sad, 2 = I am sad all the time and I can’t snap out of it, and 3 = I am so sad or unhappy that I can’t stand it). The range of the measure is 0–63 with higher scores suggesting more severe depressive symptomatology. The internal consistency was good (21 items; $\alpha = 0.81$).

Acceptance and Action Questionnaire (AAQ-2). The level of psychological flexibility was evaluated using the Acceptance and Action Questionnaire (AAQ-2; Bond et al., 2011). The AAQ-2 is a self-report measure consisting of 10 statements (e.g., I worry about not being able to control my feelings.) The statements are rated for accuracy on a seven-point scale where 1 = never true and 7 = always true. The range is 10–70. High scores indicate high psychological flexibility in our data. The internal consistency was acceptable (10 items; $\alpha = 0.66$).

Adult State Hope Scale (ASHS). Hopefulness was evaluated with

the Adult State Hope Scale (ASHS; Snyder, 2000), which consists of statements about goal-directed energy and planning to accomplish one’s goals (e.g., I can think of many ways to accomplish my goals.) There are six statements in the measure. The statements are rated for how true they feel at the moment. The scale for rating the items is 1–8 (1 = definitely false, 8 = definitely true) with the range of the total score being 6–48. High scores represent high hopefulness. The internal consistency was good (6 items; $\alpha = 0.79$).

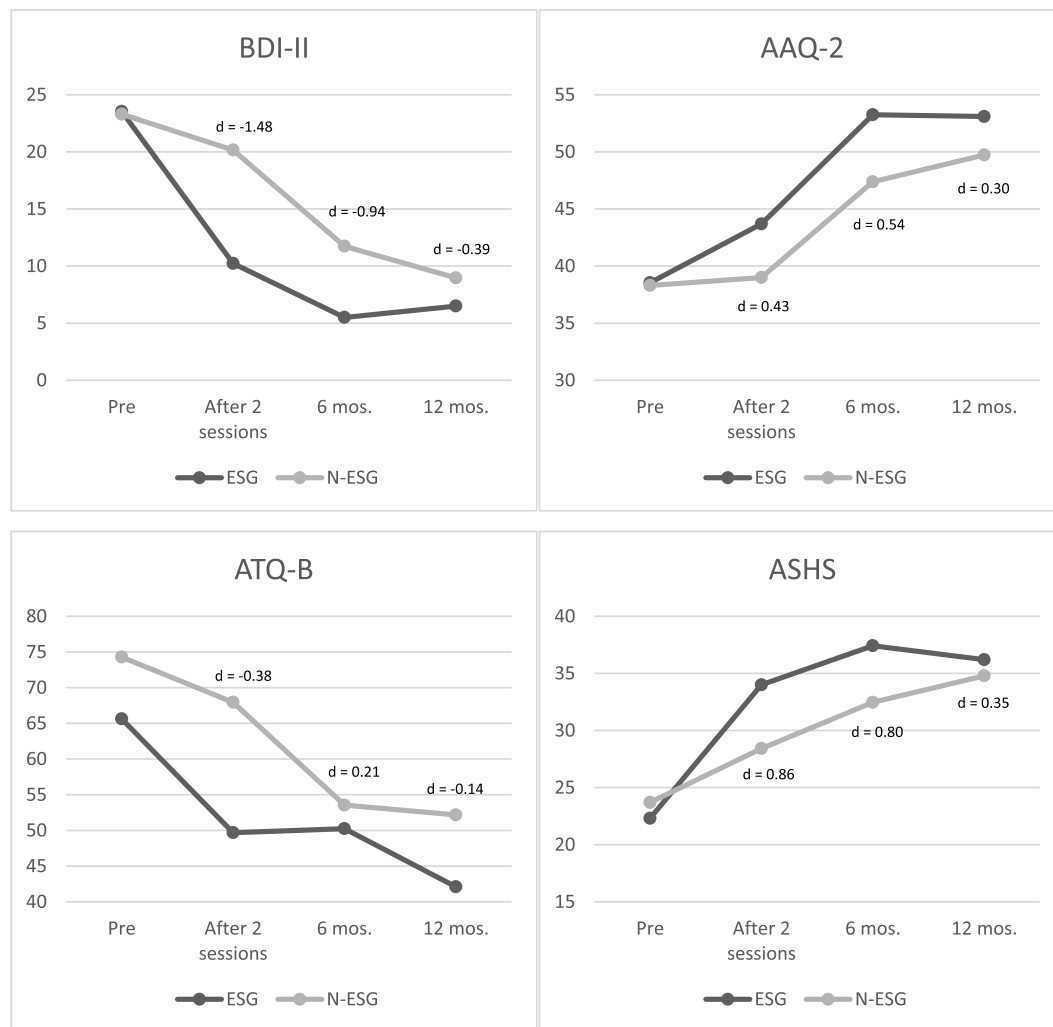
The Automatic Thoughts Questionnaire’s believability-dimension (ATQ-B). The Automatic Thoughts Questionnaire’s (ATQ; Hollon & Kendall, 1980) believability-dimension (ATQ-B) was used to evaluate the believability of depressive thoughts. The ATQ questionnaire is comprised of a list of thoughts associated with depression (e.g., I’m worthless) that rated for how frequently the thought can be recognized and for how believable it feels. The ratings for both dimensions are on a five-point scale (1 = not at all, 5 = all the time and 1 = not at all, 5 = totally, respectively). The range of each dimension of the questionnaire is 30–150 with higher scores indicating more frequent or more believable thoughts. The internal consistency for believability was excellent (ATQ-B; 30 items; $\alpha = 0.97$).

Statistical analysis. Early sudden gainers were identified using the RCI classification (Jacobson & Truax, 1991) and was based on changes after the first two sessions (BDI-II measured approximately three weeks after the pre-measurement, $M = 23.50$, $SD = 7.79$ days). The rationale for using the RCI classification was based on the two following reasons: 1) we wanted to take into account the measurement error, and 2) to address the question of clinical significance, i.e. changes required to reflect meaningful response to the treatment (for a more detailed description, see Keinonen et al., 2018). The first two weeks represent the first third of the six-week intervention and sudden gains during this time are thus defined early sudden gains as suggested by Kelly et al. (2005). We used ESG and non-ESG (N-ESG) groupings to explore the treatment results up to the 12-month follow-up. To assess the effect size of these differences, we used the corrected Cohen’s d value, as suggested by Morris (2008), which allowed differences in group sizes and pre-treatment scores between the groups to be accounted for. To analyze the stability of changes in the ESG and N-ESG groups, we used a series of ANOVA analysis to investigate differences in the level of depression, psychological flexibility, hopefulness and believability of depressive thoughts at pre-treatment, at the six-month follow-up and at the 12-month follow-up. Eleven outliers were detected in the data. The outliers were manually transferred on the distribution of each variable. The outliers were assigned to the highest/lowest value in the data for the current variable. The ANOVA analysis were then carried out using both the original data with outliers and without outliers to see if they had an effect on the interpretation of the ANOVA analysis. The outliers did not affect interpretation and statistics are reported for the raw data. The normality assumption was met based on a visual inspection of the Normal Q-Q plots except for the ATQ-B, for which normality was violated at the follow-up phases. The homogeneity of variances assumption was also violated for individual comparisons (for BDI-II, AAQ-2, and ASHS at 6-month follow-up and for ATQ-B at 12-month follow-up). The Welch ANOVA was performed for these comparisons to account for the violation of homogeneity of variances. The between-group differences were also investigated using between-group effect sizes, where an effect size of $d > 0.20$ was considered small, $d > 0.50$ was considered moderate, and $d > 0.80$ was considered large (Cohen, 1988). The correlation between ESG and follow-up treatment results are reported as indicators of the strength of association between the variables. Pearson’s chi-squared test was used to evaluate the clinical significance and prognostic value based on RCI classification, use of medication for depression, and psychological treatment after the intervention for ESG and N-ESG at the two follow-up points (6 and 12 month). The analyses were computed using IBM SPSS Statistics 24.

Table 2

Descriptive statistics for competence, adherence, and use of ACT processes during the intervention for all 37 student therapists.

	Range	Mean	SD
Overall Competence	2.00–5.00	3.31	.87
Overall Adherence	2.00–5.00	3.23	.74
Deliteralization/Defusion	1.50–4.50	2.88	.78
Willingness/Acceptance	2.00–5.00	3.19	.78
Creative Hopelessness/Workability/Control is Problem	1.00–3.50	1.53	.60
Values and Goals	1.50–4.00	2.50	.67
Committed Action	1.00–4.50	2.49	.91



Note: Corrected *d*-values are presented, where pre-differences and different sample sizes are controlled for as suggested by Morris (2008).

Fig. 2. Level of depression (BDI-II), psychological flexibility (AAQ-2), believability of depressive thoughts (ATQ-B), and hopefulness (ASHS) at pretreatment, after two sessions, 6-month follow-up, and 12-month follow-up with between-group effect sizes. Note: Corrected *d*-values are presented, where pre-differences and different sample sizes are controlled for as suggested by Morris (2008).

3. Results

The descriptive statistics for the pretreatment and follow-up measurements are reported in Fig. 2, where the participants were grouped according to the RCI classification based on changes observed after two sessions (for descriptive values, see also Table 3.). The classification identified 13 participants (23%) as the ESG group, whereas those participants who had not improved after two sessions were referred to as the N-ESG group ($n = 43$, 77%). The mean change during the two first sessions (during a period of 2–4 weeks after the pre-measurement) in BDI-II within the ESG group was 13 scores, and in the N-ESG group, it was 3 scores (Keinonen et al., 2018). There were no significant pretreatment differences between the groups on any of the measures used.

First, given the large difference in the change of depression symptoms during the two first sessions between the ESG and N-ESG groups, we were interested in how these groups changed over time after the brief intervention ended. A series of ANOVA analyses with the ESG group as the between-groups variable and level of depression, psychological flexibility, hopefulness, and believability of depressive thoughts as the dependent variables (measured at pretreatment, 6-month follow-up and 12-month follow-up). The ANOVA analyses showed that there

were no pretreatment differences between the groups, but the ESG group reported significantly lower depression ($F_{1,40} = 10.79$, $p = 0.002$), higher psychological flexibility ($F_{1,37} = 4.70$, $p = 0.037$) and higher hopefulness ($F_{1,44} = 5.72$, $p = 0.021$) at the six-month follow-up (see Table 3). In accordance with this, the between-group effect size indicated that there was a large difference ($d \geq 0.80$) in depression symptoms, a large difference ($d \geq 0.80$) in hopefulness (ASHS), a moderate ($d \geq 0.50$) difference in psychological flexibility (AAQ-2), and a small difference in believability of depressive thoughts (ATQ-B) in favor of the ESG group at 6-month follow-up.

At the 12-month follow-up, there were no significant differences between the groups on any variable. However, when investigating the effect sizes at 12-month follow-up, there were small between-group differences ($d \geq 0.20$ - < 0.50) in depression symptoms, psychological flexibility, and hopefulness (ASHS) in favor of the ESG group. The difference in believability of depressive thoughts (ATQ-B) was very small.

Overall, as we can see from Fig. 2, the changes in the ESG group were maintained during the 12-month follow-up period. The 95% confidence intervals for change scores from pretreatment to 12-month follow-up were as follows: for BDI-II, ESG (14.21, 22.40) and N-ESG

Table 3

Descriptive statistics and one-way analysis of variance for between group differences in depression, psychological flexibility, hopefulness and believability of depressive thoughts at pre-treatment, 6-month follow-up and 12-month follow-up.

Measure	Time	M (SD)		F	df	p
		ESG	N-ESG			
Depression BDI-II	Pre-treatment	23.54 (5.14)	23.30 (7.30)	0.01	1,54	.914
	After 2 sessions ^a	10.23 (3.35)	20.16 (7.82)	43.23	1,47	< .001
	6-months ^a	5.50 (4.23)	11.74 (9.03)	10.79	1,40	.002
	12-months	6.50 (5.86)	8.97 (7.71)	0.87	1,41	.357
Psychological flexibility AAQ-2	Pre-treatment	38.54 (12.45)	38.30 (9.68)	0.01	1,54	.943
	After 2 sessions	43.69 (9.12)	39.00 (9.46)	2.49	1,54	.120
	6-months ^a	53.25 (6.38)	47.38 (12.37)	4.70	1,37	.037
	12-months	53.10 (6.44)	49.73 (11.05)	0.84	1,41	.366
Hopefulness ASHS	Pre-treatment	22.31 (8.35)	23.70 (7.98)	0.30	1,54	.588
	After 2 sessions	34.00 (9.34)	28.42 (8.69)	3.98	1,54	.051
	6-months ^a	37.42 (4.46)	32.36 (10.48)	5.72	1,44	.021
	12-months	36.20 (9.92)	34.79 (10.42)	0.14	1,41	.706
Believability of depressive thoughts ATQ-B	Pre-treatment	65.62 (20.68)	74.28 (26.42)	1.18	1,54	.283
	After 2 sessions ^a	49.69 (14.16)	67.95 (22.72)	12.16	1,32	.001
	6-months	50.25 (32.26)	53.54 (24.58)	0.14	1,49	.709
	12-months ^a	42.10 (9.55)	54.18 (25.36)	3.68	1,19	.070

Note.

^a = Welch ANOVA statistic was reported to account for violation of homogeneity of variance.

(11.20, 17.11); for AAQ-2, ESG (−7.15, −24.85) and N-ESG (−6.34, −15.17); for ASHS, ESG (−4.85, −25.95) and N-ESG (−8.01, −14.84); and for ATQ-B, ESG (14.48, 43.32) and N-ESG (11.72, 30.34).

Second, we were interested in whether the change in depression symptoms during the first two sessions was associated with the observed changes from the pre-treatment to the 12-month follow-up. When analyzing the total sample, there were significant correlations between the magnitude of ESG (i.e. early changes in depression) and the long-term treatment outcome for depression (12 months $r = 0.36$, $p \leq 0.05$, $n = 43$), psychological flexibility (12 months $r = -0.33$, $p \leq 0.05$, $n = 43$), and hopefulness (12 months $r = -0.29$, $p \leq 0.05$, $n = 43$). Change in believability of depressive thoughts up to 12-month follow-up was not significantly associated with early changes in depressive symptomatology.

Third, we were interested in the clinical significance of ESGs. The RCI classification for the ESG and N-ESG groups 6 and 12 months after the treatment support the results above. The number of participants classified as recovered in the ESG group was very high at both 6- and 12-months follow-up (Table 4) supporting the stability of early changes. However, no significant differences between the groups were obtained. Also, there was no difference between the groups in the use of medication for depression during the follow-up. Very few of the participants in the ESG group received treatment during the follow-up ($n = 1/10$ at 12-month follow-up).

Table 4. RCI classification frequencies, use of medication for depression, and psychological treatment after the intervention and Pearson's chi-squared test for the early sudden gain (ESG) group and the non-early sudden gain (N-ESG) group 6 and 12 months after treatment.

4. Discussion

The aim of the current study was to explore the stability of ESGs in a six-session ACT intervention delivered by novice therapists. We have previously reported that ESG was associated with a superior treatment outcome at posttreatment (Keinonen et al., 2018). In the current study, we examined the long-term effect of ESG observed within two sessions on changes in symptom and therapeutic process measures from pre-treatment to 6- and 12-month follow-up phases.

The results suggested that ESG in depression symptoms resulted in a stable change that persisted through the follow-up phases up to 12 months. Thus, roughly 20% of clients diagnosed as having a major depressive disorder showed clinically significant changes in depression

symptoms within two sessions, and these changes were maintained 12 months after the intervention. The results were supported by the RCI classification at 6 and 12 months after treatment: a very high number of the ESG group members who participated in the follow-up phase were still improved (recovered or improved in the RCI classification) suggesting that ESG is a very stable improvement. Further, larger changes in depression symptoms during the two first sessions predicted larger changes in depression symptoms, psychological flexibility, and hopefulness at the 12-month follow-up. However, those participants not showing ESG achieved an equal positive change at 12-month follow-up after the intervention.

The results are in line with those reported by Lemmens et al. (2016) in that sudden changes during the treatment do produce better follow-up results up to a certain point (i.e. five months), but the group differences eventually fade. Similarly, in the current study, the ESG group showed a better outcome at the 6-month follow-up, particularly in depression symptom. However, at the 12-month follow-up, the differences between the ESG and N-ESG groups were small. Tang et al. (2007) also reported that sudden gains were associated with lower relapse and symptom recurrence rates. Although the differences between the ESG and N-ESG groups were non-significant in the RCI classification at follow-up in the current study, only one of the 13 participants who had experienced ESGs showed symptom recurrence. However, some studies have reported sudden gains resulting in significant differences in the treatment outcome even at 12-month follow-up phases (e.g., Bohn et al., 2013; Hedman et al., 2014), whereas the current results failed to show a significant difference after 12 months.

Although fewer studies have been published focusing on the long-term effect of ESG than those focusing on the immediate effect, these findings support the view that changes in symptomatology very early in the treatment predict stable and significant improvements. In addition, ESGs appear to indicate a very low rate of symptom recurrence even when the treatment is delivered by novice therapists. It may be important to include regular assessment in the early phase of the treatment to capture and to take advantage of ESGs.

It has been reported elsewhere, that around 60–70% of those participants who took part in the current study benefitted of the six sessions ACT intervention (Kyllönen et al., 2018). These results suggest that almost half of those participants who benefit from a six-session ACT intervention delivered by novice therapists, benefit after only two sessions. However, further research is needed to examine whether a very short treatment could be sufficient on its own for a subgroup of

Table 4

RCI classification frequencies, use of medication for depression, and psychological treatment after the intervention and Pearson's chi-squared test for the early sudden gain (ESG) group and the non-early sudden gain (N-ESG) group after two sessions and 6 and 12 months after treatment.

		ESG n = 13	N-ESG n = 43	χ^2	p
After two sessions n = 56	RCI-classification				
	Recovered	12 (85%)	–	56.00	< .001
	Improved	2 (15%)	–		
	Unchanged	–	42 (98%)		
	Deteriorated	–	1 (2%)		
6 months n = 15	Medication				
	Yes	3 (23%)	17 (40%)	1.18	.228
	No	10 (77%)	26 (60%)		
		ESG n = 12	N-ESG n = 39	χ^2	p
	RCI-classification				
12 months n = 43	Recovered	12 (100%)	23 (59%)	7.17	.067
	Improved	–	8 (21%)		
	Unchanged	–	7 (18%)		
	Deteriorated	–	1 (3%)		
	Medication				
	Yes	2 (17%)	14 (36%)	1.58	.209
	No	10 (83%)	25 (64%)		
	Psychological treatment				
	Yes	1 (8%)	6 (15%)	0.39	.535
	No	11 (92%)	33 (85%)		
		ESG n = 10	N-ESG n = 33	χ^2	p
	RCI-classification				
	Recovered	9 (90%)	24 (73%)	1.84	.398
	Improved	1 (10%)	4 (12%)		
	Unchanged	–	5 (15%)		
	Medication				
	Yes	4 (40%)	13 (39%)	0.00	.973
	No	6 (60%)	20 (61%)		
	Psychological treatment				
	Yes	1 (10%)	7 (21%)	0.64	.425
	No	9 (90%)	26 (79%)		

Note: Cutoff C = 14.49. Cutoff C was calculated using the pretreatment mean and standard deviation for all data (n = 56).

participants with depression symptoms or if it is important to continue the treatment after ESG is observed to ensure the stability of the changes.

Further research is also needed to understand if it could be possible to promote sudden gains or early sudden gains. Previous research has not been able to consistently associate sudden gains with a certain method or session content. Interestingly, high therapeutic competence and long clinical experience do not appear to be necessary for ESGs to occur, either. In fact, the limited training time and the absence of clinical experience among the novice therapists who provided the treatment in the current study did not seem to significantly reduce the frequency of sudden gains compared to the frequency reported in studies that did use more experienced therapists. Given that ESGs produce long-lasting effects, it would be highly beneficial to be able to intentionally facilitate sudden gains. Furthermore, knowledge of the processes associated with ESGs could be utilized to create more effective treatment procedures if further research were able to identify factors responsible for the early sudden gains.

The occurrence of sudden gains in various settings suggests that sudden gains are not specific to certain populations, treatments, or settings and nor are they consistently related to any of the demographic or psychological measures used across sudden gain studies. In fact, there is no strong evidence to support any of the hypotheses presented to explain sudden gains. Individual studies have associated sudden gains with certain method-related factors (e.g., cognitive changes; Tang & DeRubeis, 1999) or demographic factors (e.g., participant age; Kelly et al., 2005), but none of these results have been replicated or widely supported by other studies. At this moment, sudden gains remain unexplained, and some studies have even suggested that sudden gains could be related to some of the common features of psychological treatments (Kelly et al., 2005). Therefore, future research on sudden gains and possible factors that could explain sudden gains could have

an interesting role in furthering the understanding of crucial processes in psychological interventions.

It is important to be aware of the following limitations when generalizing the results of the current study. First, the sample that was investigated was small, and 23% of participants were lost at the follow-up phases. The participants who did not find the intervention helpful may have declined participation in the follow-up more often than those who experienced the intervention as effective, creating the possibility of a biased sample. Second, the intervention was delivered by novice therapists with only four days of training. This may cause variation in the competence levels of the therapists causing more between-subject variation than might be expected in data with experienced therapists. Third, most of the participants were women, and all were volunteers who were motivated to take part in the intervention. This limits the sample's representativeness compared to clinical populations in general. Follow-up time was relatively short (12 months). Also, the intervention was continued after ESG was observed. Thus, more studies are needed investigating the maintenance of treatment effects after ESGs without any further treatment. This knowledge would help with the utilization of the prognostic value of ESGs, and it could also help with the development of effective brief intervention models. In addition, follow-up designs that control for possible treatment during the follow-up period are needed to draw further conclusions regarding the long-term effects of ESGs.

In conclusion, the current results support the hypothesis that ESGs in an ACT intervention yield stable improvements in depression symptoms and psychological flexibility, which can be identified up to 12 months after treatment. Though preliminary, the results suggest that novice therapists can produce ESGs with very little training among approximately 20–25% of the clients. The findings add to the body of research that suggests that changes early on in an intervention can be used as a prognostic tool on an individual level.

Acknowledgements

We thank Ms. Katariina Ala-Karjanmaa, Mrs. Henna Kokkonen, Ms. Linda Elo, Ms. Riikkasisko Kirjonen and Dr. Päivi Lappalainen for their help in recruiting the participants and preparing the data. The study was supported by the Academy of Finland (project no. 140126 to Raimo Lappalainen).

References

- Aderka, I. M., Appelbaum-Namdar, E., Shafran, N., & Gilboa-Schechtman, E. (2011). Sudden gains in prolonged exposure for children and adolescents with posttraumatic stress disorder. *Journal of Consulting and Clinical Psychology*, 79(4), 441.
- Aderka, I. M., Nickerson, A., Bøe, H. J., & Hofmann, S. G. (2012). Sudden gains during psychological treatments of anxiety and depression: A meta-analysis. *Journal of Consulting and Clinical Psychology*, 80(1), 93–101.
- Beck, A. T., & Steer, R. A. (1987). *Manual for the revised Beck depression inventory*. San Antonio, TX: Psychological Corporation.
- Beck, A. T., Steer, R. A., & Brown, G. K. (2004). *Manual for the Beck depression inventory-II*. San Antonio, TX: Psychological Corporation.
- Bohn, C., Aderka, I. M., Schreiber, F., Stangier, U., & Hofmann, S. G. (2013). Sudden gains in cognitive therapy and interpersonal therapy for social anxiety disorder. *Journal of Consulting and Clinical Psychology*, 81(1), 177.
- Bond, F. W., Hayes, S. C., Baer, R. A., Carpenter, K. M., Guenole, N., Orcutt, H. K., ... Zettle, R. D. (2011). Preliminary psychometric properties of the acceptance and action questionnaire-II: A revised measure of psychological inflexibility and experiential avoidance. *Behavior Therapy*, 42(4), 676–688.
- Clerkin, E. M., Teachman, B. A., & Smith-Janik, S. B. (2008). Sudden gains in group cognitive-behavioral therapy for panic disorder. *Behaviour Research and Therapy*, 46(11), 1244–1250.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Dour, H. J., Chorpita, B. F., Lee, S., & Weisz, J. R. (2013). Sudden gains as a long-term predictor of treatment improvement among children in community mental health organizations. *Behaviour Research and Therapy*, 51(9), 564–572.
- Greenfield, M. F., Gunthert, K. C., & Haaga, D. A. (2011). Sudden gains versus gradual gains in a psychotherapy training clinic. *Journal of Clinical Psychology*, 67(1), 17–30.
- Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (1999). *Acceptance and commitment therapy*. New York: Guilford Press.
- Haynes, S. N., & O'Brien, W. H. (2000). *Principles and practice of behavioral assessment*. New York: Springer.
- Hedman, E., Lekander, M., Ljótsson, B., Lindefors, N., Rück, C., Hofmann, S. G., & Schulz, S. M. (2014). Sudden gains in internet-based cognitive behaviour therapy for severe health anxiety. *Behaviour Research and Therapy*, 54, 22–29.
- Hollon, S. D., & Kendall, P. C. (1980). Cognitive self-statements in depression: Development of an automatic thoughts questionnaire. *Cognitive Therapy and Research*, 4(4), 383–395.
- Hunnicutt-Ferguson, K., Hoxha, D., & Gollan, J. (2012). Exploring sudden gains in behavioral activation therapy for major depressive disorder. *Behaviour Research and Therapy*, 50(3), 223–230.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59(1), 12–19.
- Keller, S. M., Feeny, N. C., & Zoellner, L. A. (2014). Depression sudden gains and transient depression spikes during treatment for PTSD. *Journal of Consulting and Clinical Psychology*, 82(1), 102–111.
- Kelly, M. A., Roberts, J. E., & Ciesla, J. A. (2005). Sudden gains in cognitive behavioral treatment for depression: When do they occur and do they matter? *Behaviour Research and Therapy*, 43(6), 703–714.
- Lemmens, L. H. J. M., DeRubeis, R. J., Arntz, A., Peeters, F. P. M. L., & Huibers, M. J. H. (2016). Sudden gains in cognitive therapy and interpersonal psychotherapy for adult depression. *Behaviour Research and Therapy*, 77, 170–176.
- Masterson, C., Ekers, D., Gilbody, S., Richards, D., Toner-Clewes, B., & McMillan, D. (2014). Sudden gains in behavioural activation for depression. *Behaviour Research and Therapy*, 60, 34–38.
- Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods*, 11(2), 364–386.
- Norton, P. J., Klenck, S. C., & Barrera, T. L. (2010). Sudden gains during cognitive-behavioral group therapy for anxiety disorders. *Journal of Anxiety Disorders*, 24(8), 887–892.
- Keinonen, K., Kyllönen, H., Astikainen, P., & Lappalainen, R. (2018). Early sudden gains in an acceptance and values-based intervention: Effects on treatment outcome for depression and psychological flexibility. *Journal of Contextual Behavioral Science*, 10, 24–30.
- Kyllönen, H. M., Muotka, J., Puolakanaho, A., Astikainen, P., Keinonen, K., & Lappalainen, R. (2018). A brief acceptance and commitment therapy intervention for depression: A randomized controlled trial with 3-year follow-up for the intervention group. *Journal of Contextual Behavioral Science*, 10, 55–63.
- Plumb, J. C., & Vilardaga, R. (2010). Assessing treatment integrity in acceptance and commitment therapy: Strategies and suggestions. *International Journal of Behavioral Consultation and Therapy*, 6(3), 263–295.
- Snyder, C. R. (2000). *Handbook of hope: Theory, measures and application*. San Francisco: Academic press.
- Tang, T. Z., & DeRubeis, R. J. (1999). Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology*, 67(6), 894.
- Tang, T. Z., DeRubeis, R. J., Hollon, S. D., Amsterdam, J., & Shelton, R. (2007). Sudden gains in cognitive therapy of depression and depression relapse/recurrence. *Journal of Consulting and Clinical Psychology*, 75(3), 404–408.
- World Health Organization (1992). *ICD-10: International statistical classification of diseases and health-related problems*. Geneva: WHO.