



# International Journal of Transactional Analysis Research

Volume 2, Issue 2 – July 2011

ISSN 2218-3159

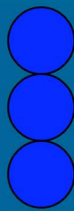
<https://doi.org/10.29044/v2i2>



## Volume 2 Issue 2 July 2011

### Contents

Editorial <i>Julie Hay</i>	2
Transactional Analysis as Psychotherapy Method – A Discourse Analytic Study <i>Roland Johnsson</i>	3
Client Assessment in Transactional Analysis – A Study of the Reliability and Validity of the Ohlsson, Björk and Johnsson Script Questionnaire <i>Roland Johnsson</i>	19
Evaluating the Outcomes of Transactional Analysis and Integrative Counselling Psychology within UK Primary Care Settings <i>Biljana van Rijn, Ciara Wild, Patricia Moran</i>	34
The Impact on Self Perception of Ego States of a Transactional Analysis Introductory Training Course (TA 101) <i>Traian Bossenmayer</i>	44



## Editorial

### Julie Hay

I continue to be intrigued and delighted with the submissions we are receiving for the journal. You will also see that I have been impressed by the way that authors are providing us with resources that are more than prompts for future researchers – much of the material can be used to enhance our self-reflection, supervision and professional practice.

This time, we have two new papers from Roland Johnsson, who, with a co-author, appeared in the first ever IJTAR issue back in July last year. He has evidently been very busy! He continues to research various aspects of therapy within Sweden.

We also have a further paper from Biljana van Rijn and Ciara Wild, both of whom also appeared in a previous issue with some co-authors – in their case the January issue this year. The current paper is co-authored with another of Biljana and Ciara's colleagues at Metanoia Institute in the UK, Patricia Moran.

Our fourth article in this issue comes from a new author, Traian Bossenmayer, a Romanian who reports on research done as part of his master's thesis undertaken at Tilburg University in the Netherlands.

So Sweden, UK and the Netherlands. And the topics are just as varied, including discourse analysis of TA psychotherapy, client assessment using a script questionnaire and checklist, the outcomes of TA (and integrative counselling psychology) therapy within primary care settings, and exploring whether attendance at TA '101' introductory courses leads to changes in self perceptions of which ego state are exhibited. There is much statistical material in this issue, and with the co-operation of the authors we have sought to maintain a balance between showing all statistical tables and showing enough to make the

papers meaningful. Those of you considering replicating their work can contact the author for more information if necessary.

We begin this issue with Johnsson's work on TA as a psychotherapy method. Building on categorisations first proposed by McNeel (1975), this study involved two assessors preparing discourse analysis of many hours of TA group therapy. Key TA components were identified and the author provides full details of the seven main and 42 sub-categories that were defined and operationalised. In addition to being the basis for further research, the list, together with Johnsson's comments about which categories seem to be rarely used, will be a great resource for reflective practice and supervision.

The second of Johnsson's articles is again a thorough piece of research, in which he and two colleagues independently assessed ten clients of a year-long therapy group, using written material from therapy termination six years ago and videotaped evidence produced six years later. To assess they used a Script Questionnaire/Checklist (Ohlsson, Björk & Johnsson 1992) and again the author has provided details, so it can be used by future researchers as well as being an invaluable resource for therapists and supervisors looking to analyse process.

Van Rijn, Wild and Moran have repeated the research methodology used as part of the UK IAPT (Improving Access to Psychological Therapies) (NHS 2011), with some additional measures. They have applied this to two groups of therapists, and demonstrated that those with a TA orientation, and those with an Integrative Counselling Psychology approach, obtain similar outcomes to those attained by CBT (Cognitive Behavioural Therapy) when working one to one over an average of nine sessions with clients referred

within primary care settings for anxiety and depression. In addition to utilising several well-known measures, the authors provide details of newly-designed adherence questionnaires relating to the two approaches under study – so yet another resource for reflective practice and supervision as well as for future researchers.

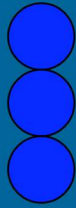
Our final paper in this issue reports on an investigation into self perceptions of ego states by participants who had attended one of three TA 101 Introductory Courses. Bossenmayer used the Adjective Check List (Gough & Heilbrun 1980) to compare participant perceptions of their own ego state behaviours at start and end of training plus one month later. Although the study suffered from falling subject numbers, there were still enough involved to produce a statistically significant change in self perceptions of Critical Parent ego state, which decreased after the training. The author has identified various limitations of the study but provides valuable information and ideas for future research. He also reminds us that we need simple models of ego states if we want clients to self-identify rather than being assessed by us as 'experts'. I am reminded of Ian Stewart's (2001) keynote speech when he used the metaphor of a filing cabinet to explain why we need different ego state models to suit different purposes.

In summary then, a pan-European suite of authors this time, with investigations into the outcomes of our work whether that be long term therapy, short term therapy, or TA introductory courses. Lots of useful ideas for future research alongside plenty of resources that we can all use to analyse our own professional processes.

I hope you will enjoy reading this issue as much as I have enjoyed choosing the articles for publication.

## References

- Gough, H.G. & Heilbrun, A.B. (1980). *The Adjective Check List Manual*. Palo Alto: Consulting Psychologist Press.
- McNeel, J.R. (1975). *A study of the effects of an intensive weekend group workshop*. Ph.D. dissertation, California School of Professional Psychology.
- NHS (2011) [www.iapt.nhs.uk/about.iapt/](http://www.iapt.nhs.uk/about.iapt/) accessed 22/7/2011
- Ohlsson, T., Björk, A. & Johnsson, R. (1992). *Transaktions-analytisk psykoterapi*. Lund: Student-litteratur.
- Stewart, I (2001) Ego States and the Theory of Theory: The Strange Case of the Little Professor *Transactional Analysis Journal*, 31 (2) 133-147



## Transactional Analysis as Psychotherapy Method – A Discourse Analytic Study

© 2011 Roland Johnsson

### Abstract

*Operational definitions of categorisations by McNeel (1975) were developed and applied by the author and an independent assessor to complete discourse analysis of 72 hours of transactional analysis group therapy in the style of Goulding & Goulding (1976, 1979) conducted during 1984/85. Results showed that the therapist used an average of 42% of the discourse space and that the therapy did indeed contain TA components, with the two main categories being 'Feeling Contact' and 'Contracts', and with particular use of TA techniques of 'talking to Parent projections', 'make feeling statement', 'mutual negotiation' and 'specificity/clarity'. Inter-rater reliability was 46.2% (Araujo & Born 1985), Cohen's (1960) kappa coefficient shows a spread from slight to moderate agreement, and the Odds Ratio (Viera, 2008) is above 1.0 for most categories.*

### Key words

transactional analysis (TA), psychotherapy, discourse analysis, TA categories, group therapy,

### Introduction

This study of TA group therapy focuses primarily on the discursive strategies, i.e. the therapist's categorised interventions. This means identifying changes and repetitions of categorised conversation processes and codings (identification) of when and how often they occur in the conversations.

### Literature review

#### Discourse Analysis

A discourse is a specific way to talk about and understand the world. It specifies the manner or pattern we use when we interact and express ourselves in different social contexts or discuss certain phenomena. The social context thus consists of what we are saying, what we accomplish with what we are saying, and what impact what is said has on us.

Discourse analysis is an analysis of emerging patterns and regularities focusing on these social exchanges of words (Foucault, 1993).

The current view that discourse is something fairly regular follows Michel Foucault (1972, 1993); he assumes a social constructionist perspective, where the truth is a discursive construction. Different knowledge regimes, such as transactional analysis, indicate what is true and false. This defines the theoretical and practical frame of reference, thereby creating conditions for the study of repetitive interventions and opportunities for categorisation and coding of therapy evidence. Discourse analysis is primarily interested in the discursive practice rather than in the individual experience.

According to discourse analysis, the client and the therapist identification are determined by the patterns that emerge here-and-now in the conversation and not by the patterns that individuals historically carry with them. It is said that the subject will become fragmented or decentred (Winther-Jorgensen & Phillips, 2000) with an increasing number of identities, depending on what discourses they are part of. The identity is changing, being represented by the position selected in the discursive context. To speak is the same as to construct an identity, according to Potter, String & Wetherell (1984).

Discursive psychology (DIP) was developed in England by Billig (1987, 1996), Edwards (1997) and Potter (1997). When applied to analysing therapy sessions and authentic conversations in different contexts it is called conversation analysis, CA (Sacks, 1992). In her research at Linköping University in Sweden, Karin Aronsson (Aronsson, 1996, 1998; Aronsson & Cederborg, 1996) focuses on 'identity-in-interaction', where the social order is an important factor. This has stimulated studies of institutional

contexts such as court trials and family therapies; she also makes analyses of 'Social Choreography', where studies of the social space (Bakhtin, 1984) in the 'communicative dance' develop in an on-going dialectical process. The positioning does not proceed from a predetermined social order but from what happens in the conversation process.

The Linköping group is interested in concepts like discourse space (Aronsson & Rundström 1988, 1989), allocation of discourse space, turn-taking control, direction of conversation, orchestration (Aronsson, 1999), allocation and definition of turns in interaction and preferential right of interpretation (Peyrot 1987; Buttny, 1990; Aronsson & Cederborg, 1996), control of choice, change and summation of topics. These different aspects on what determines communicative exchange affect balance and influence the conversation, for good and bad.

### Transactional Analysis

The creator of TA, Eric Berne, was interested in group therapy long before he developed the TA method. In some early publications (1953, 1954, 1955, 1958) he presented TA as a group therapy, exhibiting a preference for group over individual therapy because the process in the group offered a practical tool for understanding how interactions between individuals in the present moment (transactions and games) are linked to the individual and their underlying patterns (ego states and script). Berne's group therapy differs from psychoanalytically oriented groups (Bion, 1974) that see the group as a systemic, separate whole, which affects the individual's unconscious needs. One of these directions (Yalom, 1995; Rogers, 1951; Slavson, 1947, Wolf et al, 1993) emphasises the interaction between group members and the therapist as a facilitator. The psychological forces may operate freely with few therapist interventions, increasing anxiety, projection and acting out, which are then interpreted by the therapist. In Berne's group therapy, however, the therapist is an active and visible leader in every transaction (Berne 1970). This fact makes TA an adequate method to be studied with a discursive approach, where the therapist's interventions can be categorised and identified.

Discourse psychology focuses on language as social practice in interaction with others. The discursive approach differs from transactional analysis in the perception of identity stability. TA emphasises that the discursive practice, the therapy, should lead to a change by making new decisions on an emotional and cognitive level. The assumption is that, for example, a negative sense of identity has its origin in locked adaptive patterns developed in childhood, known as scripts. Both TA and discourse analysis share the

basic interactionist view, but have different views of self and identity changeability. TA emphasises instead that the exchange between people, the transaction, has a potential for change, but the underlying mental structures, as a script, limit the individual's choice and possibilities for change.

One direction in TA that developed in the 1970s was Redecision Therapy (Goulding & Goulding, 1976, 1979). Goulding's group therapy is focused on intrapsychic change in clients (rededecision). Interactions in the group are toned down in favour of *individual therapy in group* (my emphasis). The main exchange is between therapist and client, and the group acts as a resonance and support in their individual work. The group therapy in the study follows this direction.

### Categorisations used for Discourse Analysis

In a PhD dissertation by the American psychologist and TA therapist John McNeel, (1975), the major elements of this therapy were categorised. The thesis was primarily an effectiveness study. In a comparative t-test before and then three months after therapy he stated, using Shostrom's (1964) Personal Orientation Inventory (POI), that intensive therapy over a weekend (a so-called marathon) resulted in significant changes in clients in 10 of the 12 personal orientations (e.g. self-acceptance, spontaneity).

McNeel's secondary interest was to see what factors/categories of the therapy led to changes in the client. The seven main categories with their 42 sub-categories (components) form the basis for this study. These categorisations, modified and operationally defined by the author and an independent observer, are thought of as requirements of the TA method that will be met, coded and compared as a measure of TA consistency.

### Aims of the study and questions posed

The aim of this study was to examine whether the psychotherapy conducted was consistent with what the TA method requires. There is both an interest in what can *generally* be considered to describe TA and also in what is *specific* to the method.

High level of agreements (consistency) between assessors' category codings may indicate that the psychotherapy conducted follows what is generally considered to constitute a TA therapy.

Differences in coding frequency for different categories, with high correlation between the assessors and high coding frequency, may indicate what categories are specific to TA therapy.

According to the aim the following questions have been posed:

1. Can essential components of a transactional analysis group therapy be found in the study?
2. Are there agreements between codings of the independent assessor and the author in that both identify the elements constituting a TA group therapy?
3. What is the difference in agreements between the coding of sub-categories and main categories and what does this entail?
4. Do the categories describe what is defined as typical or specific to a transactional analysis group therapy?
5. What categories are in this case TA-specific?

## Ethical permission

Protocol 104-2 (Forskningsetikommittén (2002) from the Ethical Research Committee of Lund Universities meeting 20 March 2002) confirming ethical permission to use the clinical material for research.

## Methodology

### Discourse analytical study design

The following was applied to recordings of a one-year therapy group:

#### Source material

The source material consisted of 24 videotaped therapy sessions from the year 1984/85 with 10 clients and one therapist, with session lasting three hours including a coffee break. The therapist (the author) was a certified psychologist as well as a Certified Transactional Analyst (Psychotherapy) (International Transactional Analysis Association, 2004).

The therapy sessions were recorded by a sound engineer. Due to technical problems only 66 sixty-minute tapes were available out of a total of 75 therapy tapes. The transcriptions of these sessions comprised 813 pages, with an average of 65 pages per session.

A sample of 11 sessions was made so that all phases of the therapy were represented. Ten sessions made up the basic data set for the regular part of the study, Sessions 2, 4 and 6 from the beginning, 9, 11, 12 and 16 the middle and 19, 23 and 24 the end. Session 22 was used as a pilot study.

#### Clients

The clients were eight women and two men. The average age was 35 with a variation between 27 and 55. Half of them were single. Six clients had academic backgrounds. Clients were volunteers who had requested therapy at a private clinic in Malmö, Sweden

(Institute of Life Therapy – IFL), included consecutively from a waiting list. A secretary managed written and verbal information about the therapy and notifications to the group. Before the therapy began the clients were contacted via telephone by the therapist. In an individual meeting a short check of the conditions for the therapy was made. Only clients with severe disorders such as psychoses were rejected. All the first ten clients on the waiting list were accepted. Their therapy was self-funded and they had given their written consent to video-recording the therapy for research purposes.

#### The Independent Observer/Assessor

The independent assessor participated in the study from once the transcripts had been prepared. He was a 30-year-old psychology student with nothing but the written exam work left to be awarded his psychologist degree. He had no previous knowledge of TA. He coded from the transcripts without listening to the recordings, to achieve a level of blind review.

#### The pilot study

Categorisation of the pilot session revealed that the assessors had different opinions so the category contents and definitions were made more robust. We realised that the author/therapist's inside perspective and the independent assessor's outside perspective influenced the content definition of the categories. In order to achieve a good consistency in identifying the categories, but without reducing the differences in perspectives too much, we decided to begin the independent coding after our third coordination meeting. Appendix A shows the final definitions. These definitions are different from McNeel's, who used TA terminology and examples instead of definitions. All the main categories were restructured with new headings. Some categories were added, such as the main category 'Relations' with its four sub-categories. Common psychological terminology was used to define the categories.

#### The main study

The 10 sessions were then analysed. The author (A) and the independent assessor (I) coded independently. A total of 8452 codings were made; 3731 by A and 4721 by I.

#### Calculation of inter-assessor reliability

The two assessors' codings were compared and the percent agreement and kappa ratio were calculated. Full details are provided in the Results section (Appendix A).

#### The Categories and Coding Principles

The extensive pilot study was carried out before the main study to enable the assessors to obtain a mutual understanding of the meaning of the 42 categories. The assessors first used an individual interpretation procedure followed by a consensus discussion

and an agreement decision. A high number of coding options (42 categories) complicated the coding, and therefore the assessors were prepared for and trained in the use of a simplified computerised procedure ('a pop-up menu').

Two general principles for coding were formulated:

1. The therapist's statements or interventions are coded, based on the assumption that it is the therapist in interaction with the client who contributes to the therapy, following a line that is specific to the psychotherapeutic direction.

2. Each statement could be allocated to a maximum of three categories, although in most cases only one coding was used.

Statements containing "xxx" (i.e. an inaudible fragment) are excluded from coding. As the material is extensive, the loss is deemed acceptable and viewed as random.

When one, two or three codings were exactly alike, it was assessed as full agreement. When at least one of the assessors used more than one coding and this matched one coding by the other, it was considered partial agreement.

## Results

### Discourse space

Although not the main interest in the study, when reading the therapy transcripts it was noted that a communication structure emerged in which the therapist has great influence on the arrangement of the therapy session. The therapist controls the initial and final discussion, directing double-chair work, has the largest share of and influence on the discourse space as well as greatest control of choices, changes and topic summaries.

The clients' discourse space in their own therapy work is counted in the transcripts as a dialogue with the therapist. The assessors identified that the therapist

used an average of 41.7% of the total discourse space (Table 1). From the remaining 58.3%, each client's allocation of discourse space varies between 3.8 and 8.3 % of the entire therapy.

The framework and contract procedures in the TA therapy create opportunities for communication that are both controlling and permissive. The process follows a democratic dialogue methodology based on mutual negotiation, where the client's influence is supposed to be equal to that of the therapist. In the study the therapist dominates the discourse space to a fairly great extent, which might reduce the client's potential for spontaneous contributions. In the therapist role, according to TA, a combination of an active and democratic leadership with a strengthening of client power is preferable.

### Codings

#### Individual and matching codings

In Table 2 the two assessors' individual and matching codings are shown for both main and sub-categories. The difference between matching codings calculated from all the main and sub-categories also appear. Of the independent assessor's (I) 4721 and the author's (A) 3731 codings there is agreement in 1419 codings in the sub-categories and 1953 codings in the main categories, as shown in Table 3. It also shows that the agreement is generally higher (534 + codings) in the main categories, which is natural, given the more general basis. The calculated difference is large for the main categories Feeling Contact (+144) and Reality Testing (+124), but Language Usage (+93) and Strokes (+80) also show a significant difference. The lowest difference is in Contract (+18) and Relations (+5). This means that the Feeling Contact and Reality Testing categories have higher agreement (priority) in calculations made on the basis of main categories, while the Contract category has priority in the sub-category calculation. Both methods of calculation rank Language Usage the highest and Relations the lowest.

*Table 1: Discourse space for the clients and the therapist (the number of conversational turns/interventions) for each therapy session and in total as well as the therapist's percentage of discourse space*

Session	2	4	6	9	11	12	16	19	21	23	24	Total
Therapist	551	630	549	562	450	536	460	428	508	472	115	5261
Clients	629	585	676	665	530	894	833	642	1168	796	322	7740
Total number	1180	1215	1225	1227	980	1430	1293	1070	1676	1268	437	13001
Therapist % discourse space	46.6	51.8	44.8	45.6	45.33	39.5	35.6	46.2	39.2	38.3	26.1	M=41.7

Note. M = mean in percentage



*Table 2: Ranking based on the number of matching codings by main and sub-category calculation, respectively. Priority 1–7*

Priority	Main category calculation		Sub-category calculation	
	Main category	Agree	Main category	Agree
1	Language Usage	414	Language Usage	321
2	Reality Testing	369	Contract	300
3	Strokes	363	Strokes	283
4	Feeling Contact	335	Reality Testing	245
5	Contract	318	Feeling Contact	188
6	Pattern	149	Pattern	82
7	Relations	5	Relations	0
<b>Total</b>		<b>1953</b>		<b>1419</b>

Note. Agree = Agreement between the assessors

*Table 3: Ranking of the frequency of the assessors' individual and matching codings of the sub-categories. Priority 1–9*

Priority	Sub-categories	Main categories	A+I	A	I	Agree
1	Specificity/Clarity	Language Usage	1352	517	835	281
2	Mutual negotiation	Contract	994	396	598	262
3	Make feeling statement	Feeling Contact	367	158	209	100
4	Talking to Parent projections	Strokes	350	156	194	99
5	Responsibility	Reality Testing	628	237	391	95
6	Train Adult	Reality Testing	407	288	119	73
7	On the side of the Child	Strokes	482	191	291	71
8	Support/Permission	Strokes	441	161	281	62
9	Use humour	Reality Testing	329	182	147	61

Note. I = Independent assessor, A = Author, Agree = Agreement between the assessors

### *Sub-category frequency*

Some sub-categories are coded as more frequent than others, both in terms of the assessors' individual and of their jointly matching codings. A high frequency in one single assessor means that he believes that the category is commonly used in therapy. A similar high frequency agreement with the other assessor increases the reliability of one category being TA-typical. An overview of the nine most frequent assessments (Table 4) shows that the two sub-categories 'specificity/clarity' and 'mutual negotiation' are clearly the most frequent in the assessors' matchings but also in individual codings. 'Make feeling statement', 'talking to Parent projections' and 'responsibility' also have a high correspondence between assessors' matching and individual codings. However, the 'train Adult' category differs from this, as it is a highly matching coding but shows a big difference in the individual codings. A has, in relative terms, given

higher priority to this category, as compared with I. The following 15 categories may be prioritised individually or jointly, but there is no clear priority for both. The other 18 categories have low priority.

### *The assessors' individual and matching codings for the 7 main categories and their sub-categories.*

The results of the codings for each of the seven main categories and their sub-categories were reviewed (see example in Table 5). The table shows the assessors' individual and matching codings for the sub-categories in each session. The results of the main category were also reported, calculated both on the basis of the total sub-categories and on the total main categories. Moreover, it is shown in which phase of the therapy the main category is the most frequent.

Table 4: Examples of the main Contract category and 4 sub-categories with assessors' individual and matching codings

Session		2	4	6	9	11	12	16	19	23	24	Total
<b>Main categories with sub-categories</b>												
Behavioural description	I	4	0	0	0	0	0	0	0	0	0	4
	A	0	0	0	0	0	0	0	1	0	1	2
	Agree	0	0	0	0	0	0	0	0	0	0	0
Confront Parent contract	I	12	1	27	5	11	1	1	4	2	0	64
	A	0	2	0	0	0	0	0	6	0	0	8
	Agree	0	0	0	0	0	0	0	3	0	0	3
Refer to contract	I	12	4	2	7	17	17	1	1	6	3	70
	A	15	4	4	8	7	13	7	1	8	11	78
	Agree	9	2	2	3	4	8	1	0	4	2	35
Mutual negotiation	I	69	57	72	101	97	67	36	33	59	7	598
	A	54	33	39	57	38	69	28	26	41	11	396
	Agree	37	18	30	47	36	30	18	18	25	3	262
<b>Total sub-categories</b>	Agree	46	20	32	50	40	38	19	21	29	5	300
<b>Total main categories</b>	I	97	62	101	113	125	85	38	38	67	10	736
	A	69	39	43	65	45	82	35	34	49	23	484
	Agree	45	22	33	54	40	44	22	21	31	6	318

Note. I = Independent assessor, A = Author, Agree = Agreement

Table 5: The most frequent sub-categories under each main category

Main category	Most frequent sub-category
<b>Strokes</b>	Talking to Parent projections
	On the side of the Child
	Support/Permission
<b>Language Usage</b>	Specificity/clarity
<b>Pattern</b>	Expose myth and magical thinking
	Separate old scene from present impasse
<b>Reality Testing</b>	Responsibility
	Train Adult
<b>Feeling Contact</b>	Make feeling statement
	Express feelings
<b>Relations</b>	Transference

#### The main Contract category and its 4 sub-categories.

In the example in Table 4 all codings for the main Contract category are shown. It then appears that the most frequent matching sub-category is 'mutual negotiation' with 262 of a total of 300 codings, i.e. 87% of the encodings in this main category. Of all the codings in this study it represents as much as 18.5%, which makes it the second most frequent category. The other sub-categories in the table have a low frequency in the study as a whole. It should be noted that contract-related interventions are most frequent at the beginning and in the middle of the therapy, which is in Sessions 2–12.

#### Main Strokes category and its 8 sub-categories

From the codings in the Strokes category 'talking to Parent projections' is the most frequent sub-category together with 'on the side of the Child' and 'support/permission'. Together they constitute 82% of all the codings in this main category. In the study they represent a total of 16.3%. The other sub-categories in the table have a low frequency. Stroke intervention occurs mainly at the end of the therapy (Sessions 19–23).

### Main Language Usage category and its 5 sub-categories

As many as 87.5% of the codings in the main category of Language Usage derive from the 'specificity/clarity' sub-category. This is the most frequent jointly coded sub-category in the study, representing 19.8% of all categories. The other sub-categories in this main category were rare with the exception of 'word confrontation/word change'. Language Usage appears to be relatively evenly distributed throughout the therapy.

### Main Pattern category and its 5 sub-categories

Of all the codings 'expose myth and magical thinking' occurred most frequently accompanied by 'connect past scenes with present conflict'. In relation to all codings in the study, these two categories occurred relatively seldom, only 2.5 and 1.5 %, respectively. Interventions related to Pattern were more frequent at the beginning of therapy (Sessions 2–9).

### Main Reality Testing category and its 7 sub-categories

The main category of Reality Testing had three sub-categories with the highest frequency: 'responsibility', 'train Adult' and 'use humour'. In relation to all the sub-categories in the study they represent together 16 %. The other four sub-categories are insignificant. Interventions related to this main category are common in the middle of the therapy (sessions 6–19).

### Main Feeling Contact category and its 9 sub-categories

The coding rate for the main category of Feeling Contact with its nine sub-categories had high rates for the groups 'make feeling statement' and 'express feelings'. In this main category these two sub-categories had 69.6% of all codings. In the study they represent 9.2% of all codings agreed upon. In the therapy process this kind of intervention occurs evenly throughout the sessions with a slight increase towards the end (Sessions 19–23).

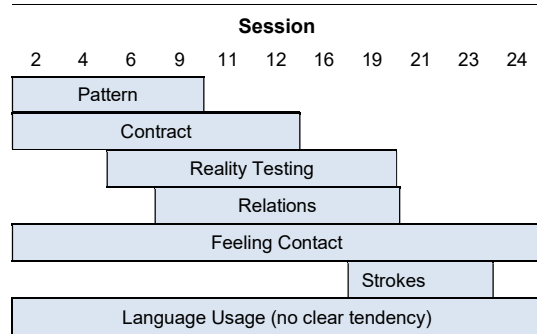
### Main Relations category and its 4 sub-categories

The last main category, Relations, had no consistent codings at all. The few existing ones had been coded as 'transference' by both examiners. I had also coded 'alliance rupture' on 43 occasions when A did not code that category at all. This coding was most frequent in the middle of the therapy process but may be considered of minor importance compared to all the categories included in the study.

### Summary of the most frequent sub-categories under each main category

A summary of the most frequent sub-categories under each main category is presented in Table 6. Compared with the ranking of the most frequent sub-categories (Table 4) it shows that the Contract interventions 'mutual negotiation' and 'specificity/ clarity' are used the most, while the Pattern interventions 'expose myth and magical thinking' and 'separate old scene from present impasse' are used the least.

Table 6: The most frequent codings of the different main categories during different phases of therapy



### Summary of the most frequent categories in different phases of therapy

A summary of the spread of codings over the different phases of therapy provides the results in Table 6. The therapy process broadly follows what is indicated in the transactional analysis literature (Berne, 1966, Goulding & Goulding, 1979; Ohlsson, Birch & Johnson, 1992; Hewitt, 1995). The Contract phase is most important in the beginning to create the alliance and the goals of therapy ('mutual negotiation'). Then comes a clarification phase where the level of awareness of the general Pattern (script) is raised and the Pattern (script) processed (the regressive phase). The client's own 'responsibility' for the problem increases in the middle of the therapy (Erskine, 1975). In the termination phase, changes are anchored with support and encouragement from the therapist. 'Feeling Contact' (Johnson & Stenlund, 2010) and 'specificity/clarity' are key elements throughout the therapy.

### The assessors' agreements

Table 7 shows that the assessors mainly used one code for each therapist intervention even if there was some diversity between them (A = 92.6%, I = 73.6%).

Table 7: Percentage distribution of the number of assessor codings for each therapist intervention, at first, second and third codings

Session	Number of ratings/interventions					
	First rating		Second rating		Third rating	
	A	I	A	I	A	I
2	91.0	82.0	7.9	15.5	1.1	2.5
4	93.0	70.8	6.5	19.3	0.5	9.9
6	94.6	71.5	5.4	17.5	0.0	11.0
9	96.6	71.7	3.4	20.6	0.0	7.7
11	96.5	64.2	3.5	21.6	0.0	14.2
12	98.5	72.8	1.5	19.2	0.0	8.0
16	88.7	77.8	11.3	16.8	0.0	5.4
19	88.7	69.7	10.8	18.4	0.4	11.9
23	90.4	70.6	8.6	17.8	1.0	11.6
24	87.8	80.3	12.2	16.0	0.0	3.7
Main	92.6	73.1	7.1	18.3	0.3	8.6

Note. I = Independent assessor, A = Author

Assessor I needed more often (+18.3%) than A (+7.1%) to encode two categories. This is understandable, given that A has an inside perspective and is familiar with the material. A third coding was used less frequently.

*Percentage of agreement for sub- and main categories*

Inter-assessor reliability was calculated partly on the basis of sub-categories and partly by main categories. Marques & McCall (2005) consider that different reliability measurements of assessor accordance create stability in qualitative research. The inter-assessment reliability percentage was calculated using a formula from Araujo & Born (1985), supplemented with Cohen's (1960) kappa coefficient and Viera's (2008) Odds Ratio. Of all 1419 assessors' matching codings (full + partial) full compatibility was coded 795 times and partial

agreement 624 times. The individual codings are relatively even for all sessions except Session 24 with its distinctly low number of codings. The mean of the percentage agreement (full + partial) was estimated at 33.5% with a relatively even distribution across all 10 sessions. When the matches from the sub-categories were recalculated to the main categories the number of consistent assessments increased by 534 to a total of 1,953. The percentage then increased to 46.2%. The calculations followed similar trends to the sub-categories.

*Complementary calculations of agreements for all sub- and main categories*

All the kappa coefficients ( $\kappa$ ) and Odds Ratios (OR) were calculated as a complement to the average percentage agreement (%), as shown in Tables 8a and 8b.

*Table 8a: Calculation of kappa quotient (k) and OR quotients (OR) for 4 main and 22 sub-categories*

Main categories	Discordant			kappa	OR
	A	I	A+I		
<b>Contract: Total calculation of the main category</b>	166	418	3323	0.44	15.23
Behavioural description	2	4	4219	0	0.00
Confront Parent contract	5	61	4156	0.08	40.88
Refer to contracts	43	35	4112	0.46	95.63
Mutual negotiation	134	336	3493	0.47	20.33
<b>Total agreement of 4 sub-categories</b>				<b>0.25</b>	
<b>Strokes: Total calculation of the main category</b>	366	597	2899	0.29	4.82
Change self-harassment to a positive fantasy	0	21	4204	0	NE
Not laughing at gallows humour	1	4	4220	0	0.00
Careful use of "Will you..?"	25	60	4124	0.26	43.99
Repetition of positive Strokes	4	1	4220	0	0.00
Support/Permission	99	219	3845	0.24	11.00
Talking to Parent projections	57	95	3974	0.55	72.65
Stroking strength and health	140	57	3993	0.24	17.51
On the side of the Child	120	220	3814	0.25	10.26
<b>Total agreement of 8 sub-categories</b>				<b>0.19</b>	
<b>Language Usage: Total calculation of the main category</b>	218	696	2897	0.35	7.90
Active use of TA terminology	3	10	4204	0.55	1121.07
Question – Re-question	2	56	4166	0.03	37.20
Hearing literally	7	71	4145	0.05	16.68
Word confrontation/word change	63	98	4035	0.26	18.95
Specificity/clarity	236	554	3154	0.31	6.78
<b>Total agreement of 5 sub-categories</b>				<b>0.24</b>	
<b>Pattern: Total calculation of the main category</b>	278	213	3585	0.31	9.02
Use fantasy	43	10	4159	0.32	125.74
Expose myth and magical thinking	104	60	4025	0.29	23.22
Separate self from others	19	36	4168	0.06	12.19
Separate old scene from present impasse	169	29	4004	0.17	18.79
Game analysis	10	145	4062	0.09	22.41
<b>Total agreement of the 5 subcategories</b>				<b>0.19</b>	

Note. Discordant = not in agreement, NE = not estimated, OR = Odds Ratio, kappa ( $\kappa$ ) = Cohen's kappa coefficient, A = Author, I = Independent assessor

Table 8b: Calculation of kappa quotient (k) and OR quotients (OR) for 3 main categories and 20 sub-categories

Main categories	Number Discordant			kappa	OR
	A	I	A+I		
<b>Reality Testing: Total calculation of the main category</b>	534	385	2937	0.31	5.27
Use intuition	7	16	4202	NE	0.00
Train Adult	215	46	3891	0.33	28.72
Responsibility	142	296	3692	0.25	8.34
Own personal power	162	43	4006	0.10	8.05
Own projections	8	20	4195	0.12	52.44
Use of video	3	2	4220	NE	0.00
Use humour	121	86	3957	0.35	23.20
<b>Total agreement of the 7 sub-categories</b>				0.17	
<b>Feeling Contact: Total calculation of the main category</b>	162	382	3346	0.48	18.11
Make feeling statement	58	109	3958	0.52	62.61
Make feeling comments	80	28	4111	0.09	11.01
Express feelings	69	125	4000	0.22	14.38
Distinction between thinking/feeling	6	11	4203	0.37	318.40
Confront 'racket feelings'	14	204	3992	0.11	20.97
Discrepancies in body language	27	1	4186	0.44	1705.41
Double-chair work	45	27	4137	0.30	54.48
Use bataca	5	7	4210	0.33	360.86
Use present tense	5	17	4202	0.08	49.44
<b>Total agreement of the 9 sub-categories</b>				0.27	
<b>Relations: Total calculation of the main category</b>	52	77	4091	0.06	5.11
Transference	56	32	4137	NE	0.00
Counter transference	1	7	4217	NE	0.00
Alliance rupture	0	43	4182	NE	NE
Boundary violation	0	0	4225	NE	NE
<b>Total agreement of the 4 sub-categories</b>					
<b>Total calculation of the 7 main categories</b>	NE	NE	NE	0.32	NE
<b>Total agreement of the 42 sub-categories</b>	NE	NE	NE	0.21	NE

Note: Discordant = not in agreement, NE = not estimated, OR = Odds Ratio, kappa ( $\kappa$ ) = Cohen's kappa coefficient, A = Author, I = Independent assessor

Kappa coefficient –  $\kappa$  – is a statistical measure of concordance, which, compared to the percentage agreement between two assessors, also takes into account accordance that occurs randomly.  $\kappa$  compares the expected consistency with the observed one, and thus gives a correction of the random factor. Norman & Streiner (2003) and Landis & Koch(1977) have described Cohen's guidelines for interpreting  $\kappa$  with 0.81 to 1.00 indicating 'almost perfect agreement', 0.61 to 0.80 'substantial agreement', 0.41 to 0.60 'moderate agreement', 0.20 to 0.40 'slight agreement' and <0.20 'poor agreement'.

Based on the main categories, four quotas have 'slight agreement', while one had 'poor', and two had 'moderate agreement'. The distribution in the sub-

categories was 15 quotas each in 'poor' and 'slight agreement', and six quotas with 'moderate agreement'.

OR is a standardised measure of effect that indicates the odds or the chance that agreement between assessors I and A in the coded category is more likely than disagreement. The quota (ratio) indicates the possibility that they agree versus the possibility that they do not agree. Odds Ratios above 1.0 strengthen the connection (association) between the assessors' matching codings and thus the probability that the assessed category is present. The tables show that all categories except six (which have 0.0) and five that could not be calculated have an OR that is above 1. The average percentage agreement broadly follows the kappa and OR values.

### Ranking of agreements for main categories

Ranking the main categories (Table 9) enables the study of the main categories which are the highest on all measurements. Feeling Contact and Contract are at the top and both have 'moderate' agreement. These sub-categories also have high coding frequencies from both assessors. Language Usage has a high frequency but slightly lower reliability.

### Ranking of sub-category agreement

Of the 21 categories, six have 'moderate' agreement (Table 10). They are 'talking to Parent projections', 'active use of TA terminology', 'make feeling statement', 'mutual negotiation', 'refer to contract' and 'discrepancy in body language'. The other 15 categories have slight agreement.

### Comparison of frequency and agreement of sub-categories

If one weighs up the sub-categories with the highest coding frequencies and reliability, the following categories are specifically important and also specific

for TA: 'mutual negotiation', 'making feeling statement', 'talking to Parent projections' and 'specificity/clarity'.

Table 9: Ranking of main category agreement according to kappa ( $\kappa$ ) and OR values, as well as percentage agreement (%)

Priority	Main category	kappa ( $\kappa$ )	%	OR
1	Feeling Contact	0.48	55.1	18.11
2	Contract	0.44	52.1	15.23
3	Language Usage	0.35	47.5	7.90
4	Pattern	0.31	44.5	9.02
5	Reality Testing	0.31	44.6	5.27
6	Strokes	0.29	43.0	4.82
7	Relations	0.06	7.2	5.11

Note: Based on k values, priority 1–2 = 'moderate' agreement, 3–6 = 'slight', 7 = 'poor' agreement. Mean kappa = 0.32

Table 10: Ranking of 21 subcategories' agreement according to the values of kappa ( $\kappa$ ), OR, and percentage agreement (%)

Priority	Subcategory	Kappa ( $\kappa$ )	%	OR
1	talking to parent projections	0.55	56.5	72.65
2	active use of TA-terminology	0.55	53.3	1 121.07
3	make feeling statement	0.52	54.6	62.61
4	mutual negotiation	0.47	52.7	20.33
5	refer to the contract	0.46	47.2	95.63
6	discrepancy in body language	0.44	44.0	1 705.41
7	distinction between thinking/feeling	0.37	35.7	318.41
8	use humour	0.35	36.9	23.20
9	use "batacka"	0.33	33.3	360.86
10	train "Adult"	0.33	35.9	28.72
11	use fantasy	0.32	32.5	125.74
12	specificity/clearness	0.31	41.5	6.78
13	double-chair work	0.30	30.7	54.48
14	expose myth and magical thinking	0.29	30.5	23.22
15	Word confrontation/word change	0.26	26.3	18.95
16	careful use of "Will you..?"	0.26	27.5	43.99
17	on the side of the "Child"	0.25	29.5	10.26
18	responsibility	0.25	30.2	8.34
19	supportive/permissive	0.24	28.1	11.00
20	stroking strength and health	0.24	26.3	17.51
21	express feelings	0.22	24.2	14.38

Note: Priority 1-6 = moderate agreement, whereas 7-21 = slight agreement based on Kappa coefficient. OR = Odds Ratio, Kappa ( $\kappa$ ) = Cohen's kappa coefficient, % = percentage agreement.

## Discussion and Conclusions

The issues are related to the overall aim to investigate whether the psychotherapy the assessors have analysed is in accordance with what is considered to be transactional analysis group therapy. The first three questions are linked to a general conclusion about what constitutes TA therapy, while the remaining two are concerned with the specificity of the method. In order to study this issue a modified discourse analytic approach was applied, where high overall agreement between the assessors' category codings was supposed to show that the psychotherapy conducted follows what is considered as constituting transactional analysis psychotherapy.

The categorisation, which acted as the assessors' coding key, was based on McNeel's thesis (1975), which was revised to create operational definitions with a general psychological content. In the pilot study the assessors made an experimental control of how well the instrument worked, and afterwards it was found that the validity of the updated coding key was good. In the ordinary study the assessors then used the coding key in order to examine how consistent and reliable the encodings could become and whether the two assessors could achieve the same results regardless of who carried out the measurement.

The main question was whether they could agree that the categories they considered to describe TA in practical work could be observed in the transcribed sessions. The answer was a calculated agreement of 33.5 and 46.2% based on the sub- and main categories, respectively. The first calculation involving the 42 sub-categories had higher precision and richness of detail than the seven generalised main categories, and gave a deeper understanding of the TA therapy components. Since sub-categories are included as aspects of the main categories, the latter reliability measure of 46.2% and a kappa coefficient of 0.32 indicate that these should still mainly describe what the assessors jointly considered to be transactional analysis group therapy.

Among the results should also be mentioned that the 20 sub-categories and two main categories (Relations and Pattern) were not coded at all or very little. The assessors had mostly one code for each intervention, which underpins the stability of agreement. That assessors on the whole used only half of the categories may be due to the difficulty of distinguishing and using a relatively large number of categories. It can of course also mean that these were representative of the therapy. A future study with codings based on the more frequent categories might provide better evidence for this.

The therapist's adherence to a method forms an important part of the result, because interventions are linked to a categorised method. The crucial issue is how purely the therapist manages to stick to 'official' theory. Canestri (2006) argues that there is a possibility that therapists develop, through further education, practical applications and personal experience, a 'private' application of the 'official' method. From the perspective of a methodological appraisal this is a problem. Nevertheless, it may be assumed as likely that the 'official' method forms the background to any new developments that can be observed and identified. Analytical (inferential) statistics have been used to make a correlational analysis of the assessors' agreements (inter-assessment reliability).

Primarily, the percentage agreement has been specified, but kappa coefficients and Odds Ratio were also calculated in order to compensate for the randomness. The significance of the measures will depend on how well the assessors can apply the previously agreed coding alternatives. The training of assessors may have led to forced consensus, which reduces their independence and thus threatens the validity of the coding categories. The validity was based on previous studies (McNeel, 1975) in which different categories were induced from an observed TA therapy. These categories were improved by operational definitions and practical evaluations of the application. The subsequent coding and data collection were thus linked to the chosen problem and research questions as well as the outcome, which by Holme & Solvang (1997) is considered essential for validity.

The specific conclusion is linked to differences in the coding rate for different categories. Categories with high reliability and a high individual frequency in the assessments show which ones are typical of or specific to TA therapy. Two sub-categories are clearly the most frequent, namely 'specificity/clarity' and 'mutual negotiation'. They are included in the main categories of Language Usage and Contract. The first main category is also individually and jointly the most common one coded with the calculation based on main categories, while the Contract category acquires much less emphasis in the more general assessment. The 'responsibility' and 'train Adult' categories, which both belong to the main Reality Testing category, were coded frequently by both evaluators and have also received many joint markers. 'Make feeling statement' obtains a great deal of agreement in the coding, even though no assessor has coded it individually to the same extent. The main Strokes category is the third most frequent and has high representation within the sub-categories 'talking to Parent projections', 'on the side of the Child' and 'support/permission'.

A number of sub-categories have been clearly emphasised in different degrees by the author and the independent assessor. 'Connect past scenes to present impasse' is marked more frequently by A, while I has coded 'confront racket feelings', and 'game analysis' more frequently. This is probably due to different perceptions of the content of the categories, since these three are transactional analysis knowledge categories, where individual knowledge and experiences have gained greater importance.

In conclusion, one can assume that the nine most frequent sub-categories show the TA categories that are most likely specific to transactional analysis psychotherapy. The question, however, remains whether these categories can also be found in other therapies and can be excluded because they may be assessed as being non-specific or 'common factors'. The therapeutic alliance is usually mentioned in this connection along with the therapist's acceptance, understanding, rational explanations and encouragement. Holmqvist (2006) and Lundh (2006) have discussed the difficulty in psychotherapy effect research of distinguishing the characteristic theory-related ingredients from common and temporary ones. Messer & Wampold (2002) as well as Luborsky et al (2002) showed that the differences between methods were small and that many 'psychotherapy interventions' are shared by most therapies. The TA method has an integrated or eclectic approach, which complicates making a clear distinction from other therapies.

Although a great many therapeutic techniques and approaches are shared, they may be practised in a way that is specific to the therapy form. Since this is not an effect study, I will confine myself to discussing what may be specific to TA, regardless of whether it is effective or not. Starting from the operational definitions, one can see that the most frequent and reliable category, 'specificity/clarity', is available in all therapies. Another highly frequent and reliable category, 'make feeling statement', can be regarded as a recurrent element in most therapies. However, 'mutual negotiation', which is often coded in agreement by the assessors, is considered to be TA-specific. TA is a contractual therapy form where mutual negotiation is an important ingredient in therapeutic cooperation.

The idea of a contract is also available in cognitive behavioural therapy (Beck, 1976, 1995) but does not permeate this form of therapy and the therapist's approach as profoundly as in TA. Another equally preferred category is 'talking to Parent projections'. It was coded in the so-called double-chair work, which is a technique originating from Gestalt therapy but is developed as a special technique in TA therapy. The therapist is schooled in this specific TA direction, named

Redecision Therapy (Goulding & Goulding, 1975), which is a likely explanation of the category's high priority.

With 'slight' (Landis & Koch (1977) reliability the two assessors have agreed about having observed transactional analysis psychotherapy in a group. This means that a description of transactional analysis psychotherapy in general terms could be made.

A large number of categories were coded a little or not at all, while a few were coded a great deal by the two assessors. Among the most frequently coded, 'mutual negotiation' is considered to be the most specific category in the TA method.

In the light of all the TA concepts and techniques that are highlighted in the study TA practitioners will find scientific support in their application of TA.

**Roland Johnsson**, lic. psychologist, lic. Psychotherapist, MSc, MSW, MA, Teaching & Supervising Transactional Analyst (Psychotherapy) can be contacted on [roland@livsterapi.se](mailto:roland@livsterapi.se)

## References

- Araujo, R. & Born, D. G. (1985). Calculating percentage agreement correctly but writing its Formula Incorrectly, *The Behaviour Analyst*, 2 (8), 207–208.
- Aronsson, K. (1996). Towards a social choreography of institutional discourse: On Pediatric encounters and children's roleplay. Plenary talk at the meeting of the American Association of Applied Linguistics, Chicago.
- Aronsson, K. (1998). Identity-in-interaction and social choreography. *Research on Language and Social Interaction*, 31, 75–89.
- Aronsson, K. (1999). Review of Family dinner talk by Shoshana Blum-Kulka. *Journal of Pragmatics*, 31, 287–292.
- Aronsson, K. & Cederborg, A-C. (1996). Coming of age in family therapy talk Perspective setting in multiparty problem formulations. *Discourse Processes*, 21, 191–212.
- Aronsson, K. & Rundström, B. (1988). Child discourse and parental control in pediatric consultations. *Text*, 8, 159–189.
- Aronsson, K. & Rundström, B. (1989). Cats, dogs and sweets in the clinical negotiation of Reality: On politeness and coherence in pediatric discourse. *Language in Society*, 18, 483–504.
- Bakhtin, M. (1984). *Problems of Dostoevsky's Poetics*. Manchester: Manchester University Press.



- Beck, A. (1976). *Cognitive Therapy and the Emotional Disorders*. New York: New American Library.
- Beck, A. (1995). *Cognitive Therapy: Basic and Beyond*. New York: The Guilford Press.
- Berne, E. (1953). Principles of Group Psychotherapy. *Indian Journal of Neurology & Psychiatry*, 4, 119–137.
- Berne, E. (1954). The natural history of a spontaneous therapy group. *International Journal of Group Psychotherapy*, 4, 74–85.
- Berne, E. (1955). Group attendance: Clinical and theoretical considerations. *International Journal of Group Psychotherapy*, 5, 392–403.
- Berne, E. (1958). Transactional Analysis: A new and effective method of group therapy. *The American Journal of Psychotherapy*, 12, 735–753.
- Berne, E. (1961). *Transactional Analysis in Psychotherapy*. New York: Grove Press.
- Berne, E. (1963). *The Structure and Dynamics of Organizations and Groups*. New York: Grove Press.
- Berne, E. (1966). *Principles of Group Treatment*. New York: Grove Press.
- Berne, E. (1970). Eric Berne as Group Therapist. *Transactional Analysis Bulletin*, 9, 75–83.
- Berne, E. (1972). *What Do You Say After You Say Hello?* New York: Grove Press.
- Billig, M. (1987, 1996). *Arguing and Thinking: A Rhetorical Approach to Social Psychology* (2nd ed.). Cambridge: Cambridge University Press.
- Bion, W. R. (1975). *Gruppterapi*. Stockholm: Prisma.
- Buttny, R. (1990). Blame-account sequences in therapy: The negotiation of relational Meanings. *Semiotica*, 78, 219–247.
- Canestri, J. (2006). *Psychoanalysis: From Practice to Theory*. London: Wiley.
- Cohen, J. A. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37–46.
- Crossman, P. (1966). Permission and Protection. *Transactional Analysis Bulletin*, 5 (19), 152–154.
- Cullberg Weston, M. (2005). *Ditt inre centrum*. Falun: Natur & Kultur.
- Edwards, G. (1997). *Discourse and Cognition*. London: Sage.
- Foucault, Michel (1972). *Vetandets arkeologi*. Staffanstorps: Cavefors.
- Foucault, Michel (1993). *Diskursens ordning*, Stehag: Symposion.
- Forskningssetikkommittén (2002). *Protokoll 104–2* from the Ethical Research Committee of Lund Universities meeting 20 March 2002
- Goulding, R. & Goulding, M. (1976). Injunctions, Decisions and Redecisions. *Transactional Analysis Journal*, 6 (1), 41–48.
- Goulding, R. & Goulding, M. (1979). *Changing Lives through Redecision Therapy*. New York: Grove Press.
- Hewitt, G. (1995). Cycles of psychotherapy. *Transactional Analysis Journal*, 25 (3), 200–207.
- Holme, I. M. & Solvang, B. K. (1997). *Forskningsmetodik*, Studentlitteratur.
- Holmqvist, R. (2006). Psykologi och placebo. *Psykologtidningen*, 10.
- International Transactional Analysis Association (2004). *Training and Exam Handbook section 7*, [http://ta-trainingandcertification.net/exam\\_handbook\\_pdfs/Section7.pdf](http://ta-trainingandcertification.net/exam_handbook_pdfs/Section7.pdf) accessed 09/06/2011.
- Jung, C. (1978). *Människan och hennes symboler*. Stockholm: Forum.
- Landis, J. R. and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174.
- Luborsky, L., Rosenthal, R., Diguier, L., Andrusyna, T. P., Berman, J. S., Levitt, J. T., et al (2002). The dodo bird verdict is alive and well-mostly. *Clinical Psychology: Science and Practice*, 9, 2–12.
- Lundh, L-G. (2006). Metaperspektiv på Psykoterapin del I & del II. *Psykologtidningen*, 6–7.
- McNeel, J.R. (1975). *A study of the effects of an intensive weekend group workshop*. Ph.D. dissertation, California School of Professional Psychology.
- Marques, J. F. & McCall, C. (2005). The application of interrater reliability as a solidification instrument in a phenomenological study. *The Qualitative Report*, 10 (3), 439–462.
- Norman, G. R. & Streiner, D. L. (2003). *PDQ Statistics* (3rd ed.). Toronto: B. C. Decker.

- Ohlsson, T., Björk, A. & Johnsson, R. (1992). *Transaktionsanalytisk psykoterapi*. Lund: Studentlitteratur.
- Peyrot, M. (1987). Circumspection in psychotherapy: Structures and strategies of counselor–client interaction. *Semiotica*, 65, 249–268.
- Potter, J. (1997). Discourse Analysis as a Way of Analysing Naturally Occurring Talk. In D. Silverman (Ed.), *Qualitative Research: Theory, Method and Practices*, 144–60. London: Sage.
- Potter, J., Wetherell, M. & Stringer, P. (1984). *Social Texts and Context: Literature and Social Psychology*. London; Routledge and Kegan Paul.
- Rogers, Carl. (1951). *Client-Centered Therapy*. Boston: Houghton Mifflin.
- Sacks, H. (1992). *Lectures on conversation: Vols I & II*. Basil Blackwell: Oxford.
- Shostrom, E. L. (1964). An inventory for the measurement of self-actualization. *Educational and Psychological Measurement*, 24, 207–218.
- Slavson, S. R. (1947). *The Practice of Group Therapy*. New York: International Universities Press.
- Viera A. J. (2008). Odds ratios and risk ratios: what's the difference and why does it matter? *Southern Medical Journal*, 101 (7), 730–4.
- Winther-Jørgensen M. & Phillips L. (2000). *Diskursanalys som teori och metod*. Lund: Studentlitteratur.
- Wolf, A., Kutash I. & Nattland, C. (1993). *The Primacy of the Individual in Psychoanalysis in Groups*. Northvale, NJ and London: Jason Aronson Inc.
- Yalom, I. D. (1995). *The Theory and Practice of Group Psychotherapy* (4th ed.). New York: Basic Books.

## Appendix A: Operational definitions of the study's seven main categories and 42 sub-categories

### I. Contract

**The client and therapist mention, quote and/or negotiate treatment contracts in some form.**

1. 'Mutual negotiation'  
The therapist starts a contract-related negotiation or responds to a negotiation initiated by the client.
2. 'Behavioural description'  
The therapist defines and substantiates a contract in behavioural terms.
3. 'Confront Parent' contract  
The therapist confronts the communication from clients in which they express their goals from a Parent position instead of listening to their own natural needs.
4. 'Refer to contracts'  
The therapist refers to the original written treatment contract or a daily contract.

### II. Strokes

**The therapist draws attention to a statement which testifies to the client's resources or confronts a self-devaluating statement. The therapist requests the client's active stance.**

5. 'Stroking strength and health'  
The therapist draws attention to new salutogenic behaviours and emotions in the client.
6. 'Repetition of positive strokes'  
The therapist repeats a positive assessment of the client, since it seems not to have been understood.
7. 'Change self-harassment to a positive fantasy'  
The therapist invites the client to replace self-torture with an enjoyable and positive imagination.
8. 'Careful use of "Will you?"'  
The therapist asks, "Will you ...?" in order to help clients to actively make their own decisions regarding a behaviour or a life situation.
9. 'Not laughing at gallows humour'.  
The therapist recognises and confronts a self-devaluating statement from the client disguised as humour.
10. 'Talking to Parent projections'  
The therapist speaks during double-chair work with the client while the client is playing the role of mother or father, as though the client were the parent at that present moment.
11. 'Support/permission'  
The therapist expresses himself non-judgmentally and encouragingly to help the client dare to express forbidden feelings and thoughts.

12. 'On the side of the Child'  
The therapist supports the client unconditionally in an attempt to express the needs, hopes and disappointments directed at authority figures from childhood.

### III. Language Usage

**The therapist asks for or makes a clarification or reformulation in terms of the here-and-now.**

13. 'Hearing literally'  
The therapist repeats a statement from the client which expresses destructive beliefs.
14. 'Specificity/clarity'  
The therapist offers or requests clarification when the client's testimony is perceived as unclear.
15. 'Word confrontation/word change'  
The therapist confronts a formulation and requests or proposes a new formulation where responsibilities are clarified.
16. 'Question–Re-question'  
The therapist repeats a question after not having received any response.
17. 'Active use of TA terminology'  
The therapist's statement contains TA terminology.

### IV. Pattern

**The therapist questions contamination or confusion, or helps the client to formulate a connection between the client's history and the here-and-now situation.**

18. 'Separate self from others'  
The therapist challenges the client to create a self-image as separate and autonomous rather than inseparably paired with someone else.
19. 'Separate old scene from present impasse'  
The therapist helps the client to distinguish how current conflict situations reflect similar scenes from childhood.
20. 'Expose myth and magical thinking'  
The therapist points out to the client unconscious and early established notions, which continue to control the client in an inappropriate way.
21. 'Use fantasy'  
The therapist invites the client to use fantasies and metaphors to playfully get an emotional image of self and own practices.
22. 'Game analysis'  
The therapist makes clear to the client in TA terms the destructive social processes the client chooses to follow.

## V. Reality Testing

**The therapist challenges clients to examine a belief about themselves, others or the world.**

23. 'Use of intuition'  
The therapist uses inspiration or an intuitive notion as a hypothesis from which clients can explore their actions.
24. 'Train Adult'  
The therapist invites the client to reflect upon and evaluate information and identify options for action.
25. 'Responsibility'  
The therapist invites clients to accept and take the consequences of the ability to affect their lives.
26. 'Own personal power'  
The therapist invites clients to accept the importance of their own choices to achieve a specific goal.
27. 'Own projections'  
The therapist invites clients to take in statements on a personal level which refer to something outside of them (e.g. "What a nice day!" to "I look nice").
28. 'Use of video'  
The therapist plays a video clip to enable the client to hear and see what took place during the therapy.
29. 'Use humour (distancing)'  
The therapist uses humour to create a distance to a subject or a situation, which is of advantage for the therapeutic process.

## VI Feeling Contact

**The therapist makes clients aware of the emotional content in client communications. The therapist stimulates and makes room for living out emotions.**

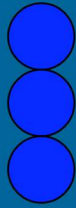
30. 'Make feeling statement'  
The therapist invites clients to express themselves verbally about their emotional state.
31. 'Make feeling comments'  
The therapist comments on the client's state of mind.
32. 'Express feelings'  
The therapist invites clients to express and show their feelings.
33. 'Distinction Between feeling/thinking'  
The therapist makes clear to clients that a feeling was asked for but a thought received in response.

34. 'Confront the racket feeling'  
The therapist confronts clients when they fall back on habitual negative emotional expressions rather than allowing themselves genuine underlying feelings.
35. 'Discrepancies in body language'  
The therapist invites clients to pay attention to the incongruence between what is said in words and what is expressed non-verbally and then asks them to express themselves congruently.
36. 'Double-chair work'  
The therapist invites the client to do what is called double-chair work. (The client improvises under the therapist's guidance a real or imagined situation usually taken from the client's history, where childhood authority figures are included and where the client may act in all of the roles.
37. 'Use bataca'  
The therapist invites the client to use a padded bat (bataca) to stimulate contact with and living out of anger.
38. 'Use present tense'  
The therapist stimulates clients to a more intensive feeling contact by encouraging them to use the present tense in descriptions.

## VII. Relations

**The communication is disturbed or interrupted by here-and-now-inadequate responses by one of the people involved.**

39. 'Transference' (crossed Parent–Child transactions from client)  
The therapist confronts clients when they express feelings and beliefs towards the therapist, which originate in their relationship to authority figures from childhood.
40. 'Counter transference' (crossed Parent–Child transaction from therapist)  
The therapist expresses feelings and beliefs towards the client which belong to the therapist's own relationship to authority figures from childhood.
41. 'Alliance rupture' (crossed transaction on an Adult–Adult transaction)  
The therapist fails to pick up and respond to the client's direct or indirect appeal for help.
42. 'Boundary violation'  
The therapist or the client goes beyond the limits agreed for the therapy.



## Client Assessment in Transactional Analysis – A Study of the Reliability and Validity of the Ohlsson, Björk and Johnsson Script Questionnaire

© 2011 Roland Johnsson

### Abstract

*A script questionnaire and associated checklist developed by Ohlsson, Johnsson & Björk (1992) was used by the author and two professional colleagues to independently assess ten clients of a year-long transactional analysis therapy group conducted by the author. Ratings based on written responses at start of therapy were compared to ratings based on videotape interviews conducted by the author six years after termination of therapy. Moderately high inter-assessor reliability was found but intra-assessor reliability was low for the independent assessors; agreement increased for script components 'primary injunction from father,' 'racket feeling,' 'escape hatch,' 'driver from father' and 'driver from mother'.*

### Key words

script, script analysis, script interview, script questionnaire, reliability, transactional analysis (TA), transactional analysis group therapy.

*Editor's Note: to allow for easier reading of the text of this paper, we have grouped all tables and figures together after the References and before the Appendices.*

### Literature Review

#### The TA Concept of Script

"The ultimate goal of transactional analysis is the analysis of scripts, since the script determines the destiny and identity of the Individual" (Berne, 1958, p. 737). Berne (1961) emphasised how scenes and experiences from early family drama are played out in everyday life in a specific and concrete way, similar to theatre dramaturgy, and argued that the task of therapy is to liberate the individual from the compulsion to repeat reliving the early script-bound scenes and thus start a new independent route in life. Although he defined script essentially as an "unconscious life plan for the individual based on decisions made in early childhood" (Berne, 1966, p. 300), he was not interested in therapies with long processes of transference and countertransference to raise an awareness of unconscious material. His method was allied with the client's functioning in the present, where the focus

was mostly on processing the early message that the client could explicitly remember. He continued to develop the concept, culminating in a definition published posthumously (Berne, 1972) of script as "an on-going programme, developed in early childhood under parental influence, which directs the individual's behaviour in the most important aspects of his life" (p. 418).

Berne's approach was further developed by his colleagues and successors (English, 1972; Goulding & Goulding, 1976, 1979; Steiner, 1967; Woollams, 1973). Steiner (1967) added the script matrix as a diagram showing how the ego states of the Child are impacted upon by injunctions, counterinjunctions, drivers and programme from the ego states of the parents. Readers unfamiliar with TA concepts are referred to Tilney (1998) for a glossary.

Steiner's matrix emphasised the functional clinical usefulness as it can be used to fill in the client's messages directly into the matrix. Other diagrams by Berne (1966), Goulding & Goulding (1979) and Woollams & Brown (1978) were more detailed and aimed at clarifying the theoretical developmental psychological aspect. Following an idea by Karpman (1966), Steiner (1967) complemented his visual matrix with a checklist where other script components were listed. Berne (1972) provided a script questionnaire comprising 220 questions; this was followed by questionnaires with fewer questions from authors such as James (1977), McCormick (1971) and Holloway (1973a).

Based on different versions of script questionnaires/checklists, Ohlsson, Björk & Johnsson (1992) designed, from their clinical experience, a script questionnaire with 43 questions (Appendix A) and a script checklist (Appendix B) including a script matrix with a checklist. These have been the work material for this study. Because of the various meanings given to the word 'script', it is suggested that the term as used in this paper *refers to all of the items in this checklist* and that, ideally, when talking about a person's script, the observer is referring to the whole checklist rather than to one or a few of the items.

## Comparable theories

Tomkins (1995a), originator of affect theory, posits nine early innate biological affects that are the foundation of our motivation to survive. When the little child communicates affects, the parents modulate these to an 'acceptable' level (Nathanson, 1992). Tomkins (1995b) makes clear that affects differ from emotions and feelings; the former are biology whereas the latter are linked to historical development and are interconnected with the individual's unique thoughts and memories, for which Tomkins (1978) also uses the term script.

Like Berne, Tomkins uses concepts and metaphors from the theatre, suggesting that feelings are organised on two levels as scenes and scripts. The scene is the basic unit, where the feeling is attached to an object (person), or a theme and an event with a beginning and an end. Tomkins' script refers to guiding principles for how the scenes are organised, and thus how specific or emotional experiences will be predicted, understood and controlled. As with TA theory, scripts can be adequate or destructive.

The cognitive theory concepts of schema (Perris, 1996), and RIGS (Representation of Interactions that have been Generalised) Stern (1991) have great similarities with Tomkins' (1978) script. They are all about individual-specific structures and patterns formed in childhood, which have subsequently guided the individual through life for good or bad. One difference is that Perris emphasises cognition while Berne, Stern and Tomkins underline the emotional interaction in early relationships and the ability to create and develop an internal object world.

TA script theory can also be linked to the psychoanalytical view on neurosis as an intra-psycho conflict (Fenichel, 1945, Haak, 1982). Small children come into conflict with the environment when they are frustrated in getting their operational needs satisfied. The conflict is pushed away, becomes unconscious and then fixated as a need at the time of the conflict. When, at times of crisis later in life, the individuals want to regain their inner balance, they regress to the point of fixation. The ego resolves the conflict by creating a symbolically designed compromise formation, the neurotic symptom. This is the solution Berne called the *early decision*, which is the basis for script formation.

In a number of studies, TA has been compared with other treatment methods (Goodstein, 1971; Ohlsson, 2010; Novey, 1999; Shaskan, Moran & Moran, 1981) where the script application of TA therapy resulted in positive outcomes.

## Diagnosis

The problem with TA diagnoses is that there is no standardisation or precision in the concepts and therefore it is uncertain whether the diagnosis has relevance (validity) in relation to the treatment process. As with most therapies, TA diagnoses are not regularly tested to achieve consistency between TA and non-TA practitioners. However, the communicability to the client and the usefulness are considered satisfactory without confirmation by a research context.

Widdowson (2010) has shown that many TA therapists use the DSM-IV or ID 10 diagnostic system in addition to their TA diagnosis. ID 10 is vaguely classified, while the DSM IV has clear behavioural criteria and can serve as a symptom classification instrument. Stewart (1996) found that DSM and ID classifications are not appropriate for practitioners because of contrasting opinions of how health problems should be described and of their narrow focus on the client's symptoms. Diagnoses do not usually follow a formally structured methodology and therapists also draw their conclusions from the informal process-oriented dialogue with the client (Cornell, 2008), in which the therapist emphasises the observation of oneself, one's feelings, memories and thoughts, so-called counter-transference. (Novellino, 1984, Hargaden & Sills, 2002). The diagnosis is then used initially in a wider sense.

The psychodynamically developed OPD-2, Operationalized Psychodynamic Diagnostics (2008), has been identified as an appropriate and well-developed diagnosis instrument, well tested in a series of reliability and validity studies. It would be important for TA practitioners to link to other systematic classifications and pragmatically create congruence between the systems. The knowledge that it is possible to describe poor health in more ways is basically fruitful and can compensate for the risk that the diagnosis has a negative effect of becoming a self-fulfilling prophecy, especially for those who believe that a diagnosis always has an organic basis and a disease. An attempt to combine diagnostic descriptions based on TA and DSM has been made by Stewart & Joines (2002) including a classification of different personality adaptations. It has become widespread among TA practitioners but has not been researched in detail.

## Aims of the study and questions posed

The aim of this study was to make client assessments, using interviews with a script questionnaire, by identifying central key conflicts in accordance with TA script theory and to examine the reliability of those analyses. The TA script theory can be viewed as a methodological theory and as an intervening variable.

The following research questions were posed:

1. Is there agreement between script analyses made on two separate occasions, on the same client and made by the same assessor (intra-assessor reliability)?
2. Is there agreement between script analyses made on two separate occasions, on the same client and by different assessors (inter-assessor reliability)?

## Ethical permission

The research was conducted under the provisions of Protocol 104-2 (Forskningsetikkomittén (2002), from the Ethical Research Committee of Lund Universities meeting 20 March 2002, confirming ethical permission to use the clinical material for research.

## Methodology

The study subjects were 10 clients who had sought therapy voluntarily and attended a one-year, 24 sessions of two and a half hours TA therapy group with the author as psychotherapist. They responded to the 43 question script questionnaire and checklist (Appendices A & B) at T1 – start of therapy and T2 – six years later. At T1 they answered the written questionnaire themselves on the basis of instructions given by the author at the first session and submitted the completed questionnaires at the next therapy session. At T2, the author acted as interviewer, using the same questions and instructions as at T1. These interviews were videotaped.

The final material consisted of nine completed script questionnaires and ten videotaped script interviews. Analyses were made on both occasions by the author and by two independent assessors separately; all three were licensed psychotherapists and formally educated transactional analysts (TSTA-P Teaching and Supervising Transactional Analyst in the Psychotherapy field) with extensive experience as trainers and psychotherapists.

A total of 57 individual analyses were completed in which 26 different script components were assessed at each analysis. A series of tables are included. Assessors coded 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> drivers from five, and made choices from 12 possible injunctions (Goulding & Goulding, 1976), three potential positions on the drama triangle (Karpman, 1968), four life positions (Berne 1972) and three variants of escape hatches (Holloway, 1973b). Other components were formulated freely. Each client was described with a document that assembled all the data from the assessments on the two occasions (see example Table 1).

Based on each client's version of Table 1, versions of Table 2 were created to show reliability of inter-assessor and intra-assessor agreement. The sum-

mary of these results is shown in Table 3 and illustrated graphically in Figure 1. In order to calculate the percentage agreement, full agreement between the three assessors was scored 3, partial agreement 2, zero for no agreement, and a hyphen was used to indicate missing assessment items. The percentage agreement was calculated as a simple and direct measure of reliability with no adjustment for random agreement in the coding. This adjustment was made at a later stage (Tables 5–8) when the kappa coefficients according to Fleiss (1971) were calculated for a sample of primary script components. Tables 9–10 focus on intra-assessor reliability.

## Reliability considerations

Sources of error with humans as measuring instruments are numerous and create well known reliability problems (Armelius & Armelius, 1985). In this study these problems were addressed by using comparisons of assessments from well-trained and experienced transactional analysts (inter-assessor reliability) and assessments on several occasions (test-retest reliability or intra-assessor reliability). The complexity of the rating procedure contributed to reducing the reliability, whereas providing direct observations of the script interviews on the second assessment gave assessors access to significant phenomenological data as if they had been there.

As the therapist conducted the video interviews himself, a clear, confident and trusting situation was created for the client. The six-year interval meant results would be influenced by the client's maturity, development and possibly by other treatments; however the long gap would decrease the client's memory of previous answers given.

Therapist adherence to methodology has been linked to important positive outcomes by Luborsky et al (1985) but the TA therapy provided in this study did not follow a specific manualised treatment procedure (adherence), and the theoretical and operational definitions of script and its different components are qualitative and multidimensional. Clinical practice in TA requires a constantly modified observational process, making it more difficult to be confident of assessor reliability in statistical terms. A logical-deductive approach was used, whilst being aware of subjective and qualitative elements in the definitions and observations that were used.

## Validity considerations

Cook & Campbell (1979) discuss problems that may occur with different types of validity. The operationalisation of the theoretical definitions of the concepts is rooted in clinical practice so construct validity is complex. Content validity has never been tested empirically, but has been assessed according to face validity by the different TA therapists. The interviews and assessments indicated that the so-called face validity was good, as the validity of the motivation,

trust and knowledge of script questionnaires validity was high among interviewers and interviewees. The therapy room where the interviews took place and the direct contact between the therapist/interviewer and the client may in this context be regarded as an authentic environment with good ecological validity (Shadish, Cook & Campbell, 2002). In the video the assessors could see how the clients reacted and responded to the interview questions. This on-line validation was built into the interview dialogue and has been used in other studies such as family therapy (Gustl et al., 2007; Sundell, Hansen, Andrée-Löfholm et al, 2006).

In a mainly qualitative study, it becomes important to describe how data have been collected and processed in a systematic manner (internal validity). The script interview in the study was compiled by the assessors and used in a clinical context over a 25-year period, so may be regarded as relevant and reliable for its intended purpose.

In clinical research the 'truth' is highly linked to practical implications so we needed to take into account the therapeutic movement or process. Kvale (1987), Polkinghorne (1983) and Malterud, (1998), report communicative and pragmatic validity as two relevant criteria; these were reflected through a careful and detailed description of how the key elements of the research took place so the reader has the opportunity to consider the transferability of the approach to similar situations (external validity).

## Results

Tables 1 and 2 are presented here as examples of how results were summarised and worked with.

### Inter-assessor reliability

The summary in Table 3 indicates that there are small variations between the two occasions. At T1 the average agreement is 59% and at T2 it is 53%.

### The total script

The assessors' agreements for the analysis of each client's total script are shown in Figure 1. The difference in client assessments is at most 24% on both occasions. There is a variation in reliability of 49–73% at T1 and 41–60% at T2. The similar matching between the assessors on the two assessment occasions for each client is acceptable. The assessors do not show any significant difference in the agreement of client assessments over time.

### Individual script components

An estimation of each script component separately (Table 3) shows that the coherence of assessments of the various components is mixed. For example, the correlation at T1 varies from 0% (the specifics of Games) to 85% (Life position) and at T2 from 0% (Counterinjunction 2 from mother) to 90% (Real feeling 1).

Script components with fixed defined categories like Driver, Injunction, Game/Drama-triangle, Life position and Escape hatch, have a higher percentage coherence compared to open categories. Especially low accordance is found in the coding of specified Games and different Counterinjunctions. The open categorisation of Racket feeling and Real feeling is an exception and has relatively high accordance.

The most significant primary components (Counterinjunction 1, Driver 1, Injunction, 1) have slightly higher coherence than the secondary and tertiary ones (e.g. Counterinjunction 2, Driver 3). This is apparent in the examination of the primary components in Tables 4–7.

The agreement between the two occasions is generally lower if one considers the individual components compared with assessments of the total script.

### Primary script components

In a second examination of the material the focus was on the script components occurring in the clients that were most obvious and most evident and, thus, were first observed (Counterinjunctions, Driver, Injunction 1, etc.). These 11 primary components (Table 4) were a starting point for a new reliability calculation based on both percentage agreement and kappa ratio.

Fleiss' kappa (1971) was used, which in contrast to Cohen's kappa is a statistical reliability measure to assess inter-assessor reliability between more than two assessors. The significance of the kappa value is determined both by the strength of the kappa quotient and by the number of categories. The kappa coefficient ( $\kappa$ ) is adjusted for randomness, as opposed to the percentage agreement (%), which leads to a stronger consistency in the correlation.

The interpretation of the significance of the Fleiss kappa ratio has been made by Landis and Koch (1977). The distribution of the study's kappa quotas on the basis of their significance intervals is summarised in Table 5. A ranking of script components has been made for T1 (Table 6) and T2 (Table 7).

According to Wood (2007), in the research context there seems to be a general view that the kappa ratio should preferably be 0.60-0.70, but that in certain cases, such as psychiatric diagnoses, a value of 0.40 and above may be acceptable. Nine categories are above 0.40 at T1 and six at T2. At T1 'Injunction from father', 'Racket feeling', 'Escape hatch' and 'Drivers from father' lie between 0.62-0.72, while at T2 only 'Real feeling' and 'Game/Drama triangle' attain such values (0.66-0.69). 'Counterinjunction 1' from mother and father has a low value on both occasions (0.15 to 0.39). The largest difference in the ratio between the two sessions relates to 'Injunction from father' with a value of 0.72 or 0.29. The total average for all of the components has a kappa ratio of 0.48.



### Intra-assessor reliability

The assessors made two analyses of each client at different times. Tables 8a, 8b and Figure 2 show that the ability to make a similar script analysis for the assessors in total is 67% for one of the assessors (C) and significantly lower, 33% and 39%, respectively, for the other two (A and B). Looking at the overall agreement based on each client, differences of 25–30% are found. Client 2 had the highest accordance (63%) between the two assessments, while Client 9 had the lowest (31%). Even an examination of the specific percentage numbers gives a picture of wide variation (20–70%) in the coherence of assessor analyses on the two occasions. Overall, it can be concluded that factors related to both the client and the assessor affect the result when assessments are made with a relatively long period in between (six years).

Ranking the results of Table 8a into Table 8b shows that the assessors have maximum coherence for clients 2 and 3, and lowest coherence for client 9.

## Discussion

The aim of the study was to assess whether you can make a diagnostically reliable script analysis using a script questionnaire. This was done by examining, with the help of two interviews, the assessors' ability to agree on client assessments. The focus was partly tied to how well the assessments match for each assessor over time (intra-assessor reliability) and partly to agreement in their analyses of the clients' total scripts and the individual components of the scripts (inter-assessor reliability). With those two measures of reliability, an indication was given of how well the script analysis on the basis of script questionnaires serves as an assessment instrument.

### Intra-assessor reliability.

The results show that assessors A and B, without any detailed knowledge of the client, made different assessments on the two occasions. Assessor C, who is the therapist and author, had much higher agreement in his two perceptions of the clients' script, which indicates that a knowledge about the client may result in more consistency in analysis although it could also mean that the assessor failed to pick up on changes. In line with Ornlinsky & Howard (1986) the large discrepancy between the reliability of different client assessments may indicate that personal variables of the client and/or assessor can play a major role in the assessment.

One explanation for the relatively low coherence is that the client has changed over time. The therapy goal and ambition is to help to change the client's script. Hence, in a successful therapy the script should not be coherent over time. Conversely, responses to the script questions could become similar even if you have changed. Most of the questions are in the nature of memories of historical events and can be expected to give similar responses, regardless of the time factor. Another possible factor is that client assessment is unreliable, because of validity problems.

### Inter-assessor reliability

When we combine all assessors' script analyses at both times and compare them with each other, the result is almost acceptable in relation to the literature. The overall correlation is 56% and relatively evenly distributed for each client. Given the difficulties with assessments over time as discussed, the overall reliability is surprisingly good. One influencing factor may be that the three assessors have worked together for a long time and have created a similar frame of reference in assessing clients. This convergence is likely to also affect the assessors' assessments over time, but becomes clearer from a general context.

When the reliability of the assessors' analysis of individual script components is examined, a considerable variation in the values is found, with the fixed categories giving better coherence than the open ones. Reliability increases significantly when examining only the 11 primary script components. More than half of those have moderate to substantial agreement and, overall, this more restrictive analysis obtains a higher reliability than the analysis of all 26 components. This is not surprising in any way but shows the difficulty of increasing the level of detail in the assessments whilst making an accurate analysis. It also shows that the gap decreases when going from the specific components to the total overall script.

There seems to be a need for an official standardised diagnostic system that can increase the reliability of psychotherapy assessments. With explicit criteria it would be easier to design and evaluate instruments that facilitate problem-formulating diagnostics (ratings and structured script interviews) and treatment follow-up (contract fulfilment). Explicit criteria would also facilitate communication between researchers, psychotherapists and clients. Finally, a clear categorical system would function as a base for decisions about mutual contracts, interventions, and well thought out treatments. Hopefully, the TA method will be researched more, and on the basis of specific descriptions and evaluation measures an alternative diagnostic classification system may subsequently develop built on whichever are the latest editions at the time of DSM, ICD or OPD, [*Diagnostic & Statistical Manual, American Psychiatric Association; International Classification of Diseases, World Health Organization; Operationalized Psychodynamic Diagnosis: Manual for Disorders and Treatment Planning, OPD Task Force (Eds)*] in which the pragmatic concepts of TA become meaningful.

Overall, it can be demonstrated that the script Interview constitutes a good and reliable basis for determining, with the help of a script checklist, a general client assessment. The ability to assess individual script components is shown to be significantly more difficult.

**Roland Johnsson**, lic. Psychologist, lic. Psychotherapist, MSc, MSW, MA, Teaching & Supervising Transactional Analyst (Psychotherapy) can be contacted on [roland@livsterapi.se](mailto:roland@livsterapi.se)

## References

- Armeliuss, B-b & Armeliuss, K. (1985). *Psykoterapiforskning – en introduktion*. Kristianstad: Rabén & Sjögren.
- Berne, E. (1958). Transactional Analysis: A new and effective method of group therapy. *The American Journal of Psychotherapy*, 12, 735–743.
- Berne, E. (1961). *Transactional Analysis in Psychotherapy*. New York: Grove Press.
- Berne, E. (1966). *Principles of Group Treatment*. New York: Oxford University Press.
- Berne, E. (1972). *What Do You Say After You Say Hello?* New York: Grove Press.
- Cornell, W. F. (2008). Setting the Therapeutic Stage: The Initial Sessions. In *Explorations in Transactional Analysis: The Meech Lake Papers*. Pleasanton, CA: TA Press.
- English, F. (1972). Sleepy, spunky and spooky. *Transactional Analysis Journal*, 2, 64–73.
- Fenichel, O. (1945). *The Psychoanalytic Theory of Neurosis*. London: Routledge, Kegan & Paul.
- Flaiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76 (5), 378–382.
- Goodstein, M. (1971). *A comparison of Gestalt and Transactional Analysis therapies in marathons*. Ph.D. dissertation, Illinois Institute of Technology
- Goulding, R. & Goulding, M. (1976). Injunctions, Decisions and Redecisions. *Transactional Analysis Journal*, 6 (1), 41–48.
- Goulding, R. & Goulding, M. (1979). *Changing lives through Redecision Therapy*. New York: Grove Press.
- Gustle, L-H et al (2007). Blueprints in Sweden. Symptom load in Swedish adolescents in studies of Functional Family Therapy (FFT), Multisystemic Therapy (MST) and Multidimensional Treatment Foster Care (MTFC). *Nordic journal of psychiatry*.
- Haak, N. (1982). *Allmänneuroslära på psykoanalytisk grund*. Stockholm: Natur & Kultur.
- Hargaden, H. & Sills, C. (2002). *Transactional Analysis: a relational perspective*. Hove: Brunner-Routledge.
- Holloway, W. (1973a). *Clinical Transactional Analysis with Use of Life Script Questionnaire*. Medina: Midwest Institute For Human Understanding.
- Holloway, W. (1973b). Shut the Escape Hatch, no IV. *The monograph series of Midwest Institute for Human Understanding no I–X*. Medina: Midwest Institute For Human Understanding.
- James, M. (1977). *Techniques in Transactional Analysis for Psychotherapists and Counselors*. Menlo Park: Addison-Wesley.
- Karpman, S. (1966). Teaching Transactional Analysis in T-groups. *Transactional Analysis Bulletin*, 5 (20), 184–185.
- Karpman, S. (1968). Fairy tales and script drama analysis. *Transactional Analysis Bulletin*, 7 (26). 39–43.
- Kvale, S. (1987). Validity in the Qualitative Research Interview. *Methods*, 1 (2).
- Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174.
- Luborsky, L., McLellan, A. T., Woody, G. E., O'Brien, C. P. & Auerbach, A. (1985). Therapist success and its determinants. *Archives of General Psychiatry*, 42, 602–611.
- Malterud K. (1998). *Validitet. Kvalitativa metoder i medicinsk forskning*. Lund: Studentlitteratur.
- McCormick, P. (1971). *Guide for Use Of Life Script Questionnaire in Transactional Analysis*. Berkeley: Transactional Pubs.
- Nathanson, D. L. (1992) *Shame and Pride. Affect, Sex and the Birth of Self*. New York: W.W. Norton & Company.
- Novellino, M. (1984). Self-analysis of counter-transference. *Transactional Analysis Journal*, 14 (1) 63–75.
- Novoy, T. (1999). The effectiveness of transactional analysis. *Transactional Analysis Journal*, 29, 18–30.
- Ohlsson, T. (2010) Scientific evidence base for transactional analysis in the year 2010. Annex 2 – the Psychotherapy List: References to research on Transactional Analysis psychotherapy effects 1963–2010. *International Journal of Transactional Analysis Research*, 1 (1), 24–29.
- Ohlsson, T., Björk, A. & Johnsson, R. (1992). *Transaktionsanalytisk psykoterapi*. Lund: Studentlitteratur.
- OPD Task Force (Eds.). (2008). *Operationalized Psychodynamic Diagnosis OPD-2: Manual for Diagnosis and Treatment Planning* (2nd Revised ed.). New York: Hogrefe & Huber.
- Orlinsky, D. E. & Howard, K. I. (1986). Process and outcome in psychotherapy. In S. Garfield & A. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (3rd ed.). New York: Wiley.
- Perris, C. (1996). *Kognitiv terapi i teori och praktik*. Stockholm: Natur & Kultur.
- Polkinghorne, D. E. (1983). *Methodology for the human sciences*. Albany: SUNY Press.

- Shadish, W., Cook, T. D. & Campbell, D.T. (1979). *Quasi-experimentation: Design and Analysis Issues for Field Settings*. Boston: Houghton-Mifflin.
- Shaskan, D., Moran, W. & Moran, O. M. (1981). Influence of group psychotherapy: A thirty year follow-up. *Transactional Analysis Journal*, 11, 270–274.
- Steiner, C. (1967). A Script Checklist. *Transactional Analysis Bulletin*, 6 (22), 38–39.
- Steiner, C. (1976). *Beyond Games and Scripts by Eric Berne – Selections from his major writings*. New York: Grove Press.
- Stewart, I. (1996). *Developing transactional analysis counselling*. London: Sage.
- Stewart, I. & Joines, V. (2002). *Personality adaptations – A New Guide to Human Understanding in Psychotherapy and Counselling*. Nottingham and Chapel Hill: Lifespace Publishing.
- Sundell, K., Hansson, K., André-Löfholm, C. et al (2006). *Multisystemisk terapi för ungdomar med allvariga beteendeproblem*. Stockholm: Socialstyrelsen.
- Tilney, T. (1998). *Dictionary of Transactional Analysis*. London: Whurr Publishers.
- Tomkins, S. S. (1978). Script Theory: Differential magnification of Affects. In *Nebraska Symposium on Motivation (1979)*, 201–233. Lincoln: University of Nebraska Press.
- Tomkins, S. S. (1995a). The quest for primary motives. In E. V. Demos, (Ed.) *Exploring Affect. The Selected Writings of Silvan S. Tomkins*, 27–63. New York: Cambridge University Press.
- Tomkins, S. S. (1995b). Role of the specific affects. In E. V. Demos, (Ed.) *Exploring Affect. The selected writings of Silvan S. Tomkins*, 68–85. New York: Cambridge University Press.
- Widdowson, M. (2010). *Transactional Analysis: 100 Key Points*. London: Routledge.
- Wood, J. M. (2007, October 3). Understanding and Computing Cohen's Kappa: A Tutorial. <http://www.wpe.info/vault/wood07/Wood07.pdf> Retrieved on 12/06/2011.
- Woollams, S. (1973). Formation of the script. *Transactional Analysis Journal*, 3, 31–37.
- Woollams, S. & Brown, M. (1978). *Transactional Analysis*. Dexter: Huron Valley Press.

## Tables and Figures

Table 1: Example: Summary of assessments for a client

Client 1	Assessor A		Assessor B		Assessor C	
	T1	T2	T1	T2	T1	T2
<b>Script components</b>						
<b>Mother:</b>						
Counterinjunction 1			Work hard	Work hard	Work hard	
Counterinjunction 2			Please others	Be kind		
Driver 1	Be strong	Be strong	Please others	Be strong	Be strong	Be strong
Driver 2	Try hard				Try hard	Try hard
Injunction 1	Don't be close	Don't be close	Don't be important	Don't be close	Don't be close	Don't be close
Injunction 2		Don't be a child	Don't be a child	Don't belong	Don't belong	Don't be a child
<b>Father:</b>						
Counterinjunction 1		Be funny	Work hard	Be funny	Be funny	Be funny
Counterinjunction 2				Be happy	Be happy	
Driver 1	Try hard		Be perfect		Be perfect	Be perfect
Driver 2	Please others	Please others		Please others	Please others	Please others
Driver 3						
Injunction 1	Don't think	Don't think	Don't belong	Don't grow up	Don't belong	Don't belong
Injunction 2	Don't feel	Don't be close	Don't be close	Don't feel	Don't feel	Don't feel
<b>Early decisions</b>	To be happy and kind on the outside but hiding the inside	Be generous, funny, please others to distance your loneliness,	Lonely must take care of myself. Nobody believes in me.	Be happy and strong take care of mum's discontent and father's fear for feelings	Be strong and tough not showing feelings but suffer in silence	Please others withdraw so others can avoid feelings so not to feel lonely
<b>Racket feeling 1</b>	worried/fear	lonely	worried/fear	happy	worried/fear	lonely
<b>Racket feeling 2</b>	sadness	sadness	guilt		guilt	sadness
<b>Real feeling 1</b>	angry	angry	angry	angry	angry	angry
<b>Real feeling 2</b>			sad	sad		
<b>Game/drama triangle</b>	Rescuer	Rescuer	Victim	Rescuer	Rescuer	Rescuer
<b>Specific game</b>	I'm Only Trying To Help You	Rapo, Clown				
<b>Life position</b>	I'm not OK-You are OK	I'm not OK-You are OK	I'm not OK-You are OK	I'm not OK-You are OK	I'm not OK-You are OK	I'm not OK-You are OK
<b>Escape hatch</b>	Not open	Not open	suicide	Not open	Not open	Not open
<b>Specific</b>	addiction			Be a loner	Be a loner	Be a loner

Table 2: Examples of inter- and intra-assessor reliability for Client 1

Client 1	Intra-assessor reliability			Inter-assessor reliability		Most frequently	Frequency
	Assessors			T1	T2		
Script components	A	B	C				
Mother:							
Counterinjunction 1	-	3	0	2	0	Work hard	3
Counterinjunction 2	-	0	-	0	0	-	0
Driver 1	3	0	3	2	3	Be strong	5
Driver 2	0	0	3	2	0	Try hard	3
Injunction 1	3	0	3	2	3	Don't be close	5
Injunction 2	0	0	0	0	2	Don't grow up	3
Father:							
Counterinjunction 1	0	0	3	0	3	Be funny	4
Counterinjunction 2	-	0	0	0	0	Be happy	2
Driver 1	0	0	3	2	0	Be perfect	3
Driver 2	3	0	3	2	3	Please others	5
Injunction 1	3	0	3	2	0	Don't belong	3
Injunction 2	0	0	3	2	2	Don't feel	4
Early decision	2	0	2	2	2	-	-
Racket feeling 1	0	0	0	3	2	Worried/fear	3
Racket feeling 2	3	0	0	2	2	Sadness	3
Real feeling 1	3	3	3	3	3	Angry	6
Real feeling 2	-	3	-	0	0	Sad	2
Game/Drama triangle I	3	0	3	2	3	Rescuer	5
Specific Game	0	-	-	0	0		2
Life position	3	3	3	3	3	I'm OK-You're OK	6
Escape hatch	3	0	3	2	3	No open	5
Specifics	0	0	3	0	2	Be a loner	3

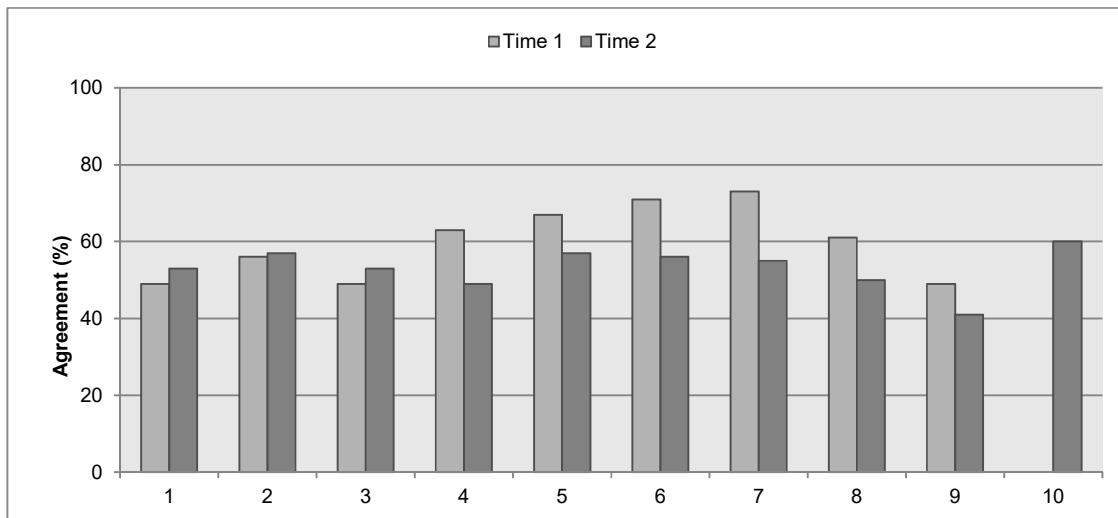
Note. 3 = full agreement, 2 = two of three agreed, 0 = no agreement, and - = no assessment. Most frequently = the most frequent component, and Frequency = number based on 0–6 possible assessments.

Table 3: Percentage of inter-assessor reliability for all clients, components and total

Client Component	T1										T2											
	1	2	3	4	5	6	7	8	9	%	1	2	3	4	5	6	7	8	9	10	%	
Mother:																						
Counterinjunction 1	2	0	2	2	3	0	-	2	0	46	0	0	0	2	2	0	0	0	0	0	0	13
Counterinjunction 2	0	2	0	-	-	-	-	-	-	22	0	0	0	-	-	0	0	-	0	-	0	0
Drivers 1	2	3	2	3	3	2	3	2	3	85	3	3	3	2	3	3	2	3	2	3	2	80
Drivers 2	2	-	2	2	-	-	0	-	2	53	0	0	0	2	0	-	3	0	3	3	3	40
Drivers 3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Injunction 1	2	0	2	2	3	2	3	2	0	59	3	2	3	2	0	2	0	3	2	3	2	66
Injunction 2	0	3	0	0	2	-	2	2	-	43	2	2	2	0	3	3	2	0	0	0	2	53
Injunction 3	-	3	-	-	-	-	2	2	-	78	-	3	-	-	-	-	3	3	-	2	3	91
Father:																						
Counterinjunction 1	0	2	0	0	-	0	-	3	0	23	3	2	3	0	-	0	0	0	0	0	-	33
Counterinjunction 2	0	0	0	-	-	-	-	2	-	17	0	2	0	-	-	-	-	0	0	-	0	13
Drivers 1	2	2	2	3	3	3	2	3	2	81	0	2	0	3	2	3	2	0	3	0	0	50
Drivers 2	2	0	2	0	-	-	0	0	0	19	3	0	3	-	0	0	0	0	0	0	-	25
Drivers 3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Injunction 1	2	3	2	3	3	3	3	0	2	78	0	2	0	3	3	2	2	2	3	2	3	63
Injunction 2	2	2	2	2	2	3	3	3	2	77	2	3	2	2	2	3	3	2	0	2	70	
Injunction 3	-	3	-	-	0	-	2	-	0	41	-	0	-	0	0	-	0	0	-	-	0	0
Early decision	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Racket feeling 1	3	0	3	3	2	3	3	3	2	81	2	2	2	0	3	2	2	2	2	2	3	66
Racket feeling 2	2	0	2	-	2	0	3	-	-	50	2	-	2	0	3	-	3	-	2	0	0	57
Real feeling 1	3	2	3	2	0	3	3	0	3	70	3	3	3	2	3	2	3	3	2	3	90	
Real feeling 2	0	-	0	2	2	-	-	2	0	33	0	-	0	0	-	-	-	3	0	-	0	20
Game	2	3	2	3	2	3	2	3	2	81	3	3	3	3	3	2	3	3	2	3	93	
Specific game	0	0	0	0	0	-	-	0	-	0	0	0	0	0	2	0	2	0	0	0	0	13
Life position	3	2	3	2	3	3	2	3	2	85	3	2	3	3	2	2	2	3	3	2	83	
Escape hatches	2	3	2	3	2	3	2	3	3	85	3	3	3	2	0	3	3	3	2	3	83	
Specific	0	2	0	-	2	2	-	2	2	48	2	2	2	2	0	2	0	2	0	3	50	
Rating	-	-	-	-	-	-	-	-	-	-	3	2	3	3	0	2	0	2	3	0	60	
<b>Total %</b>	<b>49</b>	<b>56</b>	<b>49</b>	<b>63</b>	<b>67</b>	<b>71</b>	<b>73</b>	<b>61</b>	<b>49</b>	<b>59</b>	<b>53</b>	<b>57</b>	<b>53</b>	<b>49</b>	<b>57</b>	<b>56</b>	<b>55</b>	<b>50</b>	<b>41</b>	<b>60</b>	<b>53</b>	

Note. 0 = not agreed, 2 = two of three agreed, 3 = all three agreed. - = no assessment. Calculation Example: Counterinjunctions 1 = 2 + 0 + 2 + 2 + 3 + 0 + 2 + 0 = 11. Divided with the number of assessments 5, which is multiplied by the ideal situation where everyone agrees, that is 3, which becomes 46%.

Figure 1: Graph of assessors' percentage of agreement for each client's total script at 2 times of assessments



*Table 4: kappa ratio and percentage agreement between the 3 assessors of 11 primary script components calculated for a total of 10 clients at the 2 assessment times*

Script components	T1		T2	
	kappa quota (κ)	Agree (%)	kappa quota (κ)	Agree (%)
Counterinjunction 1 from mother	*0.32	70	*0.15	50
Driver 1 from mother	**0.56	86	**0.52	86
Injunction 1 from mother	*0.43	66	**0.46	66
Counterinjunction 1 from father	*0.38	73	**0.39	66
Drivers 1 from father	**0.62	86	**0.41	70
Injunction 1 from father	**0.72	90	0.29	73
Racket feeling	**0.68	83	**0.49	66
Real feeling	*0.51	76	**0.66	90
Game/drama triangle	**0.48	86	**0.69	90
Life positioner	**0.44	86	0.36	83
Escape hatch	**0.65	90	0.38	90

Note \* p <0.01 \*\* p <0.001 Counterinjunction = primary Counterinjunction Average value is κ = 0.48 and 78%, respectively

*Table 6: Ranking of script components for 10 clients at T1 assessment*

Priority	Script components	kappa quota (κ)	Agreement (%)
1	Injunction 1 from father	**0.72	90
2	Racket feeling	**0.68	83
3	Escape hatch	**0.65	90
4	Drivers 1 from father	**0.62	86
5	Drivers 1 from mother	**0.56	86
6	Real feeling	*0.51	76
7	Game/drama triangle	**0.48	86
8	Life position	**0.44	86
9	Injunction 1 from mother	*0.43	66
10	Counterinjunction 1 from father	*0.38	73
11	Counterinjunction 1 from mother	*0.32	70

*Table 5: Distribution of kappa coefficients for 11 primary script components for 10 clients at the 2 assessment times*

Intervals of kappa coefficients	Interpretation of intervals	kappa (κ) T1	kappa (κ) T2
< 0	Poor agreement	0	0
0.0–0.20	Slight agreement	0	1
0.21–0.40	Fair agreement	2	4
0.41–0.60	Moderate agreement	5	4
0.61–0.80	Substantial agreement	4	2
0.81–1.00	Almost perfect agreement	0	0

*Table 7: Ranking of script components for 10 clients at T2 assessment*

Priority	Script components	kappa quota (κ)	Agreement (%)
1	Game/drama triangle	**0.69	90
2	Real feeling	**0.66	90
3	Driver 1 from mother	**0.52	86
4	Racket feeling	**0.49	66
5	Injunction 1 from mother	**0.46	66
6	Driver 1 from father	**0.41	70
7	Counterinjunction 1 from father	**0.39	66
8	Escape hatch	0.38	90
9	Life positioner	0.36	83
10	Injunction 1 from father	0.29	73
11	Counterinjunction 1 from mother	*0.15	50

Table 8a: Percentage of intra-assessor reliability of 9 clients at the 2 assessment times

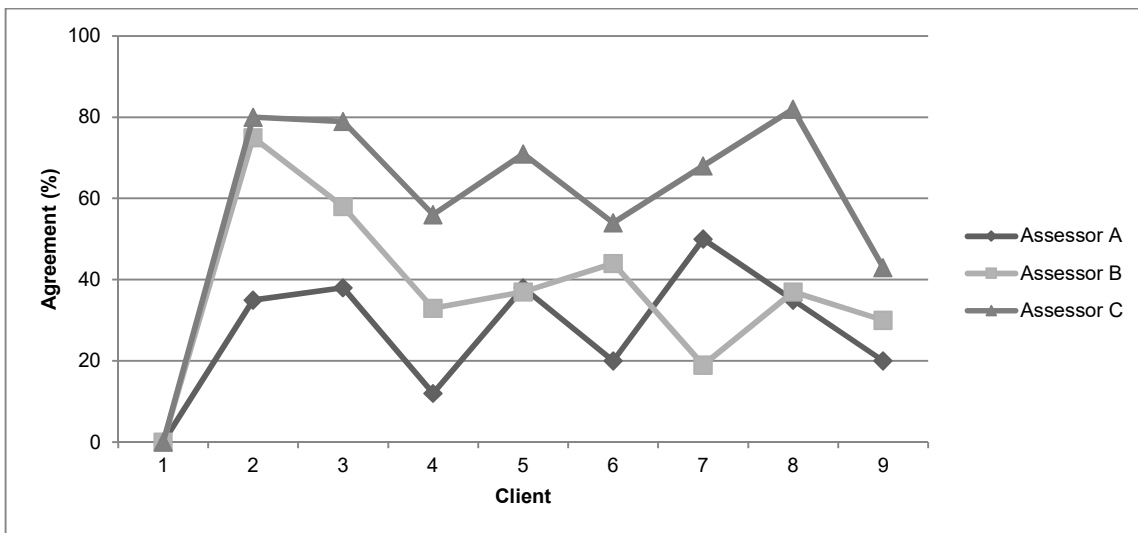
Client	Assessor A	Assessor B	Assessor C	Total (M)
1	53	20	68	47
2	35	75	80	63
3	38	58	79	58
4	12	33	56	34
5	38	37	71	49
6	20	44	54	39
7	50	19	68	46
8	35	37	82	51
9	20	30	43	31
Total (M)	33	39	67	46

Table 8b: Ranking of the percentage of intra-assessor reliability of 9 clients at the 2 assessment times

Priority	Assessor A Client	Assessor B Client	Assessor C Client	All Assessors Client
1	1	2	8	2
2	7	3	2	3
3	3	6	3	8
4	5	5	5	5
5	2	8	1	1
6	8	4	7	7
7	6	9	4	6
8	9	1	6	4
9	4	7	9	9

Note Ranking where 1 indicates maximum coherence and 9 lowest

Figure 2: Percentage agreement between script analyses made at 2 times by 3 assessors (intra-assessor reliability)





## APPENDIX A: Script questionnaire (Ohlsson, Björk & Johnsson, 1992)

1. Your name:

What does your name mean? How old are you?

2. Describe yourself briefly as you are now.

3. How do you earn your living?

4. How and with whom do you live now?

5. Do you have children? If yes, who are they, and who is the other parent?

6. What is your education and work experience?

7. Tell us what you know about your own birth?

8. Who are your biological parents?

9. Where and with whom did you live as babies?

10. As a 4–5 year old?

11. When did you begin school?

12. And when were you a teenager?

13. Describe your mother briefly, as she was when you were little?

14. Describe your father briefly, as he was when you were little?

15. What did your mother do when she was displeased with you?

16. What did she do when she was pleased with you?

17. What did your father do when he was displeased with you?

18. What did he do when he was pleased with you?

19. When you were small, what did you like best about your father?

20. As far as you remember, which is the worst memory of your father?

21. And the worst memory of your mother?

22. What did you like best about your mother?

23. What fairy tale or story did you like best as a child?

24. Tell me about the story.

25. What is it that appeals to you in the story?

26. What would you do and how did you feel when you were little and the grown-ups were stupid?

27. Do you remember any time it happened when you were little?

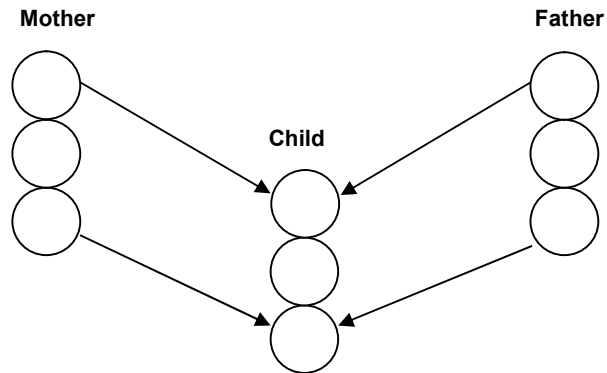
28. How do you think you are going to die? How old will you become?

29. What do you think people will say about you after your death?
30. What would 'happiness' be for you?
31. What is your biggest problem right now?
32. What is the most common bad feeling you have experienced in your life?
33. Tell me what you think your life will be like in five years.
34. What do you dislike most in yourself?
35. What do you like most about yourself?
36. What was the most important decision in your life?
37. In what way would you have liked your mother to have been different?
38. In what way would you have liked your father to have been different?
39. What kind of a life did your grandparents live? Grandmother/s? Grandfather/s?
40. Tell me which of your parents or grandparents are dead. How did they die and how old were they?  
Mother? Father? Grandmother/s? Grandfather/s?
41. If you were a magician, what would you want to conjure up in yourself?
42. What of all this do you think you can achieve even though you are not a magician?
43. Do you want to tell us anything else? Is there anything I did not ask that you think I need to know to understand your situation?

## APPENDIX B: Script analysis form

Name: \_\_\_\_\_

No: \_\_\_\_\_



Counterinjunction: \_\_\_\_\_

Counterinjunction: \_\_\_\_\_

Drivers: \_\_\_\_\_

Drivers: \_\_\_\_\_

Injunction: \_\_\_\_\_

Injunction: \_\_\_\_\_

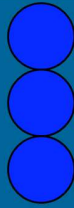
Early decisions: \_\_\_\_\_

Racket feeling/Real feeling \_\_\_\_\_

Game: (Persecutor, Rescuer, Victim) Specify: \_\_\_\_\_

Life position: \_\_\_\_\_

Escape hatch: \_\_\_\_\_



## Evaluating the Outcomes of Transactional Analysis and Integrative Counselling Psychology within UK Primary Care Settings

© 2011 Biljana van Rijn, Ciara Wild, Patricia Moran

### Abstract

*The paper reports on a naturalistic study that replicated the evaluative design associated with the UK National Health Service initiative IAPT – Improving Access to Psychological Therapies (CSIP 2008, NHS 2011), as previously used to assess Cognitive Behavioural Therapy (CBT), with the aim of evaluating 12-session treatments for anxiety and depression, applying Transactional Analysis and Integrative Counselling Psychology approaches within real clinical settings in primary care. Standard outcome measures were used in line with the IAPT model (CORE 10 and 34, GAD-7, PHQ-9), supplemented with measurement of the working alliance (WAI Horvath 1986) and an additional depression inventory BDI-II (Beck, 1996), and adherence to the therapeutic model using newly designed questionnaires. Results indicated that severity of problems was reduced using either approach, comparative to Cognitive Behavioural Therapy; that initial severity was predictive of outcome; and that working alliance increased as therapy progressed but was not directly related to outcomes. Adherence was high for both approaches. Several areas for enhancements to future research are suggested.*

### Key words

Transactional analysis psychotherapy, integrative counselling psychology, CORE, WAI, BDI-11, PHQ-9, GAD-7, anxiety/depression, IAPT, CBT.

### Introduction

The research, which took place between 2008 and 2010, with 78 clients and nine therapists, was a product of collaboration between Metanoia Institute in London and the local Primary Care Trust (PCT) in the London

Borough of Ealing. The PCT funded the project to provide and research the effectiveness of short-term Transactional Analysis (TA) and Integrative Counselling Psychology (ICP) psychotherapy in the surgeries of General Practitioners (GPs).

PCTs are those parts of the UK National Health Service (NHS) that are responsible for the local commissioning and provision of services in the first line of health care. The national context was that the UK National Institute for Health and Clinical Excellence (NICE) had issued Guidelines (still in effect) that Cognitive Behavioural Therapy (CBT) was to be the main treatment of choice for anxiety and depression. CBT was, and is, therefore, offered on a large scale within the NHS, including as part of the overall NHS initiative for Improving Access to Psychological Therapies (IAPT) (NHS, 2011).

The PCT, which already provided generic counselling in primary care, wanted to evaluate humanistic and integrative approaches within this health setting in order to make the case for offering wider therapeutic approaches within the NHS and not be limited to CBT. In order to meet that objective the project replicated the methods used within the IAPT initiative (Clark et al, 2009; CSIP, 2008).

Metanoia Institute is a TA psychotherapy training institute which also offers training in ICP and other humanistic approaches. Metanoia Counselling and Psychotherapy Service (MCPS) is the clinical service within the institute that provides low cost treatment to the public and placements for students at Metanoia Institute. MCPS has been engaged in routine clinical evaluation using CORE (CORE Information Management Systems Ltd.) for over ten years.

## Literature Review

### *Previous Research*

GP surgeries within the UK are significant in the provision of psychological treatments and the first point of contact for clients within the NHS. GP interventions include general medical care, medication, psychological intervention or a combination of these approaches, as well as referrals for medical, surgical and psychological treatments. Previous research studies have compared:

- usual GP care and counselling with effects of medication. Bedi et al (2000) found that both counselling and antidepressants were equally effective and that this was not related to patient preference. Rowland, Bower, Mellor-Clark, Heywood and Hardy (2000) conducted a systematic review of research evidence for the effectiveness of counselling in primary care and found that patients who received counselling, rather than just the usual GP care, demonstrated an improvement in symptoms coupled with a high degree of patient satisfaction.
- effects of medication and psychological therapies. (Bower, Rowland & Hardy, 2003) conducted a systematic review and meta-analysis of counselling and primary care. They found that brief counselling demonstrated significant short-term improvements in mental health and concluded that it would be a useful addition to mental health services.
- Different psychological therapies, primarily CBT, psychodynamic and interpersonal or person centred. In addition to other 'common factors' research, studies by Mellor-Clark, Connell, Barkham and Cummins (2001), Stiles, Barkham, Mellor-Clark and Connell (2008) and Stiles, Barkham, Twigg, Mellor-Clark and Cooper (2006), conducted in naturalistic settings, found that different orientations were similar in effectiveness, although there were differences in effectiveness of individual therapists.

Even though research is available on both TA and integrative approaches, these approaches have not been subject to systematic evaluation in primary care settings. In the UK this is becoming particularly important in the context of national policies, where clinical guidelines used in the health service do not recommend these treatments.

### *Transactional Analysis and Integrative Counselling Psychology*

TA, from its introduction during the 1960s by Eric Berne (Berne, 1961) has been developed, practised and taught within different professional and national contexts and evaluated in different therapeutic settings (Novey, 1999; Thunnissen, Duivenvoorden & Trijsburg, 2001; Ohlsson, 2002).

### *Integrative Counselling Psychology*

ICP as taught at Metanoia Institute is rooted in the movement for psychotherapy integration and based on the findings of the 'common factors' research (Orlinsky, Grawe, & Parks, 1994; Lambert & Ogles, 2004; Wampold, 2001). Counselling Psychology theory at Metanoia Institute focuses on building integrative frameworks for the therapeutic process, drawing on relational psychoanalysis, systemic, humanistic and existential views of the person.

### *Research Methodologies*

Evaluative research raises questions about the type of evaluation appropriate in each clinical setting, and whether it will be based on efficacy or effectiveness research design. Effectiveness research is based on clinical practice and takes place in clinical settings, but often suffers from incomplete data and inadequate monitoring of the therapists' approach. Efficacy research design is highly structured, takes place in a laboratory setting, clients and therapists are carefully chosen and the work is closely supervised. This type of research has a high degree of internal validity, but clinicians often find it hard to relate it to the realities of therapeutic practice.

Nathan, Stuart and Dolan (2000) cite the efforts to bridge the gap between the two research designs by developing effectiveness/efficacy clinics which combine features of both. This has been attempted by the National Institute of Mental Health in the US and under the IAPT initiative (Clark et al, 2009) within the UK and this project uses the same approach to evaluation.

## Aims

The aim of the project was to replicate the IAPT approach in order to explore whether TA and ICP psychotherapy, when applied in the same type of setting (GP surgeries), with similar patient groups and the same duration of treatment, would result in similar outcomes to those reported for CBT.

Additional measures were introduced to allow analysis of adherence to the model, working alliances and sessional outcomes.

## Ethical Considerations

Ethical approval for the project was given by Metanoia Institute Research Committee, an independent academic body, approved by Middlesex University.

All clients were given information about the project and signed consent forms.

## Research Design and Participants

The project was a naturalistic quantitative evaluation, which was designed to replicate the IAPT (CSIP, 2008; Clark et al, 2009) approach to evaluation. As such, it explored treatment outcomes and impact within the 'real world' of clinical practice: specifically it consisted of evaluating the effectiveness of 12-week TA and ICP therapy provided within four allocated GP surgeries.

There were no experimental or control groups, clients were referred through normal channels and were not specifically selected for the research, and there were no researcher hypotheses to prove or reject. Unlike most efficacy research, the therapists were not expected to follow a manual.

### Clients

All clients received a routine GP assessment and were referred as suitable for counselling. They were randomly assigned to TA or ICP. 78 were referred and seen for assessment, 75 attended their first treatment session, 61 (78.2%) were seen for more than six sessions and 17 (21.8%) were seen for fewer than six sessions.

Treatment was defined as the assessment session plus at least one treatment session.

The client group had a wide age range (20 to 88 years). Approximately 74% of clients were female and 38% from minority ethnic backgrounds. 36% of the clients were unemployed and only 19% were in full time employment. At assessment 77.25% of clients were above the clinical cut off for anxiety and depression, and 45.16% were above the clinical cut off on CORE 34.

### Therapists

Therapists were senior students in the third or fourth years of their training for MSc Transactional Analysis Psychotherapy or a Doctorate in Counselling Psychology that incorporates Integrative Psychotherapy. All had volunteered to take part in the project.

### Assessors

Clinical supervisors who assessed the adherence to the theoretical approach were independent senior practitioners who had ongoing supervisory relationships with the students and were accustomed to using the learning outcomes outlined in the questionnaires, albeit in a different format.

These supervisors had an assessment role at Metanoia Institute, but were not expected to assess quality of practice, which was evaluated quantitatively. Their role included offering help to therapists if they were struggling to maintain adherence to the model.

## Research Measures Used

Table 1 sets out the various measures used and at which sessions.

*Table 1: Measures used per session*

	Every session	1st session	6th session	12th session
<b>BDI-11</b>		x	x	x
<b>CORE 10</b>	x	x		
<b>CORE 34</b>		x	x	x
<b>GAD-7</b>	x	x		
<b>PHQ-9</b>	x	x		
<b>WAI</b>	From 2nd session			

### 1st session (Assessment)

At the first session each client was assessed using a battery of standard assessment measures: CORE 34, BDI-II, PHQ-9, GAD-7 and CORE 10, as described and referenced below. In addition to that, the assessment focused on establishing motivation and aims for treatment.

### Sessional measures: post each session

Sessional measures were:

- Client Health Questionnaire or PHQ-9 (Kroenke, Spitzer & Williams, 2001): 9-item questionnaire which distinguishes between clinical and non clinical populations;
- General Anxiety Measure, GAD-7 (Spitzer, Kroenke, Williams & Lowe, 2006): 7-item questionnaire which was initially developed for the Generalised Anxiety Disorder and found to have sensitivity for other anxiety disorders (Kroenke, Spitzer, Williams, Monahan & Lowe, 2007);
- CORE 10 (CORE Information Management Systems Ltd, 2007): 10-item questionnaire focusing on categories of well being, functioning, problems/symptoms and risk.
- Working Alliance Inventory, WAI (Horvath, 1986): 12-item questionnaire developed to measure working alliance as defined by Bordin (1979) – used from the second session.

### Pre, mid and post therapy measures:

Additional measures used at the first, sixth and the twelfth session were:

- BDI-II (Beck, 1996): 21-item questionnaire measuring depression
- CORE 34 (Barkham et al, 2001): 34-item questionnaire focusing on categories of well being, functioning, problems/symptoms and risk and distinguishing between clinical and non clinical populations.

The mid therapy measures aimed to collect data about the flow of therapy and increase the percentage of full data sets where the clients did not complete the full treatment.

### Adherence to the model

Adherence questionnaires were designed based on the core skills, theoretical knowledge and attitudes taught within each course, as defined in course handbooks and externally validated by Middlesex University and the national umbrella bodies of UK Council for Psychotherapy (UKCP), British Association for Counselling and Psychotherapy (BACP) and the British Psychological Society (BPS). Adherence was measured using a 5-point scale, across 16 items for TA and 24 items for ICP. Details are given in the Appendix. The ratings were done on the basis of student presentations and three audio recordings of sessions for each client. Audio recordings were familiar to therapists and supervisors as part of normal supervisory practice.

## Results

Outcomes are presented in the format which broadly follows the evaluation by Clark et al (2009), for ease of comparison. An overall data set, with comparisons between the TA and ICP groups, is followed by the statistics related to each outcome measure. Finally, there is statistical analysis of associations between measures.

### Data set

78 clients referred for therapy attended the assessment session; 38 were allocated to TA therapists and 40 to CPI therapists. Independent t-test found no significant differences in pre therapy scores on CORE 34 and BDI-11 between the two groups. 75 clients attended the first treatment session. Of these, 17 (21.8%) attended fewer than six sessions and 61 (78.2%) attended more than six sessions. The average number of sessions attended was nine, and 60% of all endings were planned.

As shown in Table 2, the sessional measures had a high percentage of completeness (97%) and allowed the analysis of outcomes at the end of therapy, even where the 'pre and post' measures were missing. The level of data completeness for CORE 34 was lower at 70.5% and BDI-II at 73.1%.

Table 2: Participants and data completeness

Measure	Number complete	Number incomplete	% complete	% incomplete
<b>Complete</b>	48	—	61.5	—
<b>Ended</b>	30	—	—	38.5
<b>BDI-II Mid/Post*</b>	57	21	73.1	26.9
<b>CORE-34 Mid/Post*</b>	55	23	70.5	29.5
<b>PHQ-9</b>	74	3	97	3
<b>GAD-7</b>	74	3	97	3
<b>CORE-10</b>	74	3	97	3
<b>WAI</b>	74	3	97	3
<b>6 Sessions +/-</b>	61	17**	78.2	21.8**

Note. \* Calculated using only clients who have completed up to session 6 or are ongoing and therefore have a mid therapy measure (if this was their last measure before they ended sessions this is also counted as their Post or final measure).

\*\* Less than six sessions completed.

### Treatment outcomes: CORE 34 and BDI-11

Table 3 shows the means for CORE 34 and BDI-11 at the three points in time. Results of a paired t-test used to investigate the differences between pre, mid and post scores showed that significant differences existed at  $P < 0.05$  between scores for the pre-post ( $t = 4.341$ ) BDI-II and mid-post BDI ( $t = 4.524$ ), mid-post ( $t = 4.606$ ) and pre-post ( $t = 4.418$ ) Core 34 totals, mid-post ( $t = 3.064$ ) and mid-post ( $t = -3.744$ ) and pre-post ( $t = -3.261$ ).

Due to unplanned endings there was a smaller number of clients (55-57%) who completed these measures before and at the end of therapy. IAPT guidelines (Clark, 2009) indicate completeness of 6% at their Doncaster site and 56% in the Newham site for the Core 34 outcome measure.

Table 3: BDI-II and CORE 34 Descriptive Statistics

Measure	Number	Mean	Std. deviation
Pre therapy BDI	78	28.7	13.3
Mid therapy BDI	57	27.2	13.7
Post therapy BDI	45	18.8	13.6
Pre CORE	77	62.5	24.6
Mid CORE	54	60.4	24.6
Post CORE	42	45.2	23.4

*Treatment outcomes: sessional measures*

Descriptive statistics for the sessional measures are shown in Table 4. A paired t-test showed a significant difference at  $P < 0.05$  between scores for pre-post PHQ-9 ( $t = 3.233$ ), pre-post GAD-7 ( $t = 4.842$ ), mid-post CORE 10 ( $t = 3.064$ ) and pre-post CORE 10 ( $t = 3.877$ ).

The high percentage of completed data sets for final session outcomes, at 97%, is well above the IAPT guideline of 80% (D.M. Clark et al, 2009).

Table 5 shows the improvement status for client from the first assessment to the end of therapy. Therapy is defined as a minimum of an assessment plus one session.

Overall, mean scores decreased towards the end of therapy but not in a linear direction. The standard deviation of scores increased and decreased across sessions and between participants showing periods of greater variation. This suggested that some clients became more distressed before getting better.

To ascertain further the meaning of outcomes, severity of the problems was analysed for all clients at the beginning and the end of therapy. The percentage of clients recorded as non clinical across the sessional measures was 43.36% (between 33.3% and 49.4%) reduced from 77.25% at the start of therapy. Non clinical was defined using the same parameters as the IAPT report (Clark et al, 2009) where scores of nine or over on the PHQ-9 are considered clinical and scores of

seven or over on the GAD-7 are considered clinical. Scores of ten or above on the CORE 10 are considered clinical following the Core 10 manual.

The average percentage of clients who have improved using sessional measures was 57.7% (between 55.1% and 64.1%).

*Working Alliance Inventory (WAI)*

The working alliance was regarded as an essential factor to the effective therapeutic process, as taught within both training programmes and supported by common factors research (Norcross, 2002). Results in Table 6 show that the mean for the working alliance increased as the therapy progressed. A paired t-test showed that there was a significant difference between mid-post ( $t = -3.744$ ) and pre-post ( $t = -3.261$ ).

*Table 6: Working Alliance Inventory descriptive statistics*

	N	Min	Max	Mean	Std. deviation
<b>Session 2 total</b>	51	33	84	60.49	12.383
<b>Session 6 total</b>	51	34	84	65.92	12.818
<b>Session 11 total</b>	38	38	84	68.53	13.506

*Table 4: Sessional measures PDQ-9, GAD-7, CORE 10 descriptive statistics*

	Pre therapy			Mid therapy			Post therapy		
	Mean	SD	N	Mean	SD	N	Mean	SD	N
<b>PHQ-9</b>	14.1	7.3	75	11.5	7.3	56	8.9	6.2	39
<b>GAD-7</b>	12.7	6.2	76	10.1	6.6	53	8.0	5.8	39
<b>CORE 10</b>	19.5	7.8	77	17.7	8.9	56	13.9	8.0	37

*Table 5: Improvement status PHQ-9, GAD-7 and Core 10*

	PHQ-9		GAD-7		CORE 10	
	N	%	N	%	N	%
<b>Improve</b>	44	54.5	47	60.3	50	64.1
<b>No change</b>	11	14.1	16	20.5	14	17.9
<b>Deteriorate</b>	21	26.9	15	19.2	14	17.9
<b>No date</b>	2	2.6				
<b>Total</b>	78	100	78	100	78	100



### *Associations between variables*

A Chi Squared was carried out to examine possible associations between variables and showed:

- significant associations ( $p < 0.01$ ) between completion status and improvement on the BDI-11
- significant associations in the completion of more than six sessions and improvement on the BDI-11, PHQ-9, GAD-7 and CORE 34. A regression was employed to ascertain whether client severity at the start of therapy predicted scores at the end of therapy. It was found that first session scores significantly predicted ( $P < 0.05$ ) final session scores on all of the measures.

Repeated ANOVA showed no significant differences between theoretical orientation and improvement on any of the measures, nor between individual therapists and improvement on any of the measures.

### *Adherence to the model*

Results for adherence to therapeutic model suggested that practitioners' adherence to their models were similar and high, 86% adherence for TA and 81% for ICP psychotherapy.

## Discussion

### *The Results*

IAPT services within the UK specialise in treatment of anxiety and depression and the IAPT report (Clark et al, 2009) found that on entering the service, approximately 86% of clients were scoring above the clinical cut off on the depression and anxiety measures being used. Generic primary care counselling, as in this project, receives referrals assessed for a 'lower intensity treatment' (CSIP, 2008) that is not based on individual diagnostic categories. However, 77.25% of the clients seen in this research project were also classified as above the clinical cut off for anxiety and depression at the start of therapy. These results suggested that depression and anxiety were more widespread in clinical populations and supports (Drozd & Goldfried, 1996) critique that single diagnostic categories may not be the best basis for evaluation or signposting for treatments.

Treatment outcomes showed that TA and ICP psychotherapy achieved change for an average of 57.7% of the clients referred, comparable to the IAPT demonstration sites (Clark et al, 2009) using CBT, which showed a change for 55% clients who have attended at least twice. Although the figures were limited by a small sample, the high percentage of full data sets for sessional measures suggested robustness in comparison to IAPT evaluation. Clark, Fairburn & Wessely (2008) and Clark et al (2009) demonstrated data completeness for sessional measures of 88.3%

and 99.6% in their two demonstration sites, in comparison to 97% within this research project. IAPT data completeness for CORE 34 was 6% and 56% in the two demonstration sites. Data completeness for CORE-34 in this research project of 70.3% was far higher.

These results showed that brief TA and ICP psychotherapy were comparable in their effectiveness to CBT within a primary care setting, when measured in a similar way.

Completing a number of questionnaires each session may have had an impact on the outcomes. Lambert et al (2002) and Miller, Duncan, Brown, Sorrel and Chalk (2006) found that clients used evaluation measures as one of the ways to give feedback to the therapist on their experiences. This feedback improved outcomes, such as achievement of reliable change and better attendance. Follow up information might clarify whether the gains were maintained in the longer term, but this was not available in this setting.

### *The Therapists and the Orientation*

There were no differences shown in effectiveness between the therapists or the orientations even though the therapists showed a high level of adherence to treatment models. This was expected on the basis of the common factors research (Smith & Glass, 1977; Asay & Lambert, 1999; Wampold, 2001; Lambert & Ogles, 2004) and the Dodo bird effect (Rozenweig, 1936).

However, the expectations of difference in the performance of individual therapists (Mellor-Clark, Connell, Barkham, & Cummins, 2001; M. J. Lambert & Ogles, 2004) was not met. The therapists worked with similar clients and numbers and all performed to a steady level. This may have been related to the similarity in their training background or to being coached by supervisors within the same training environment, where the different approaches share the same overall philosophy. Outcomes could be different within different training establishments.

The adherence questionnaires were experienced as long and sometimes cumbersome by both therapists and supervisors. They are a new measure, which needs to be developed further and standardised.

Using supervisors as assessors was experienced as helpful by the practitioners as it became part of the supervision process and offered structure and focus. However, despite their professional role and assessment experience, existing relationships with supervisors might have led to bias. The impact of the coaching was also not quantified. For greater fidelity, use of research supervisors, or random checks of recordings by the researchers, could be recommended for future projects, with and without accompanying coaching.

## The Working Alliance

Working alliance outcomes showed that the alliance increased within the duration of therapy. Therapists found the measure useful to attend to potential ruptures in the relationship, but there was no evidence that the strength of the working alliance predicted the outcomes, even though it was expected that this would be a factor (Horvath & Bedi, 2002). However, clients who stayed in therapy longer had better outcomes, so the strength of the working alliance may have had an impact on the clients' ability to use therapy.

It may be, of course that completing the WAI had an impact on the therapy itself.

## Methodology

The research team was able to use the IAPT (CISP 2008) methodology to evaluate the effectiveness of treatments in real clinical settings and achieve sufficient data completeness to enable them to reach their conclusions.

For a more complete evaluation, baseline measurements prior to assessment and follow up measures would have been useful: it would have been particularly useful to evaluate whether clients were able to maintain their gains at the end of therapy.

Measuring the adherence to the theoretical model was another area for development. These new measures need to be tested in future projects, developed and standardised.

## Future Developments

The outcomes of this research point to the need to conduct further studies with larger samples. This setting did not allow for an extension of the length of therapy according to the needs of the client, or for a follow up. The outcomes showed that severity of the scores at the outset predicted severity at the end. This suggested that longer treatment may be more appropriate for clients with more severe difficulties. To address this Metanoia Institute has applied this research clinic model to its internal service, where a larger scale project started in September 2010. The new project allows for a larger sample, longer treatments, an opportunity of a follow up and comparisons between several theoretical approaches.

**Dr Biljana van Rijn** Teaching and Supervising Transactional Analyst (Psychotherapy) and the Head of Clinical Studies at Metanoia Institute. She can be contacted on [Biljana.vanRijn@metanoia.ac.uk](mailto:Biljana.vanRijn@metanoia.ac.uk)

**Ciara Wild** MSc Forensic Psychology is Researcher/Clinic Administrator at Metanoia Institute.

**Dr Patricia Moran** Chartered Psychologist in the Integrative Department at Metanoia Institute, London, where she tutors on the Doctorate in Counselling Psychology and Psychotherapy by Professional Studies

## References

- Asay, T. P. & Lambert, M. J. (1999). The Empirical Case for the Common Factors in Therapy: quantitative findings. In M. Hubble, B. Duncan, L. & S. D. Miller (Eds.), *The Heart and Soul of Change: What Works in Therapy*, 531–537. Washington, D.C: American Psychological Association.
- Barkham, M., Margison, F., Leach, C., Lucock, M., J., M.-C. & Evans, C. (2001). Service Profiling and Outcomes Benchmarking Using the CORE-OM: Toward Practice-Based Evidence in the Psychological Therapies. *Journal of Consulting and Clinical Psychology*, 69, 184–196.
- Beck, A. (1996). United States of America Patent No. 0154018392.
- Bedi, N., Chivlers, C., Churchill, R., Dewey, M., Duggan, C., Fielding, K. et al (2000). Assessing effectiveness of treatment of depression in primary care: Partially randomised preference trial. *British Journal of Psychiatry*, October 2000 (177), 312–318.
- Bordin, E. S. (1979). The Generalizability of the Psychoanalytic Concept of the Working Alliance. *Psychotherapy: Theory, Research, Practice*, 16 (3), 252–260.
- Bower, P., Rowland, N. & Hardy, G. (2003). The clinical effectiveness of counselling in primary care: a systematic review and meta-analysis. *Psychological Medicine*, 33 (2), 203–215.
- Clark, D. M., Layard, R., Smithies, R., Richards, D. A., Suckling, R., & Wright, B. (2009). Improving access to psychological therapy: Initial evaluation of two demonstration sites. *Behaviour Research and Therapy*, 47, 910–920.
- CORE Information Management Systems Ltd. Outcome Measure 34. from <http://www.coreims.co.uk> accessed 29/7/11.
- CORE Information Management Systems Ltd. (2007). CORE Net. from <http://www.coreims-online.co.uk> accessed 29/7/11.
- CSIP. (2008). Improving Access to Psychological Therapies (IAPT) Commissioning Toolkit. <http://ebookbrowse.com/iapt-commissioning-toolkit-pdf-d48273738> accessed 27/07/11
- Drozdz, J. F. & Goldfried, M. R. (1996). A Critical Evaluation Of The State-Of-The-Art In Psychotherapy Outcome Research. *Psychotherapy: Theory, Research, Practice, Training*, 33 (2), 171–180.
- Horvath, A. O. (1986). *Another Approach to Assess the Counselling Relationship: The Working Alliance*. Paper presented at the American Educational Research Association.
- Horvath, A. O. & Bedi, R. P. (2002). The Alliance. In J. C. Norcross (Ed.), *Psychotherapy Relationships That Work. Therapist Contributions and Responsiveness to Patients*, 37–69. New York: Oxford University Press.

- Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a Brief Depression Severity Measure. *Journal of General and Internal Medicine*, 16, 606–613.
- Kroenke, K., Spitzer, R. L., Williams, J., Monahan, P. O. & Lowe, B. (2007). Anxiety Disorders in Primary Care: Prevalence, Impairment, Comorbidity, and Detection. *Annals of Internal Medicine*, 146, 317–325.
- Lambert, M. J., Whipple, J. L., Vermeersch, D. A. D., Smart, W., Hawkins, E. J., Nielsen, S. L. et al (2002). Enhancing Psychotherapy Outcomes via Providing Feedback on Client Progress: a Replication. *Clinical Psychology & Psychotherapy*, 9 (2), 91–103.
- Lambert, M. J. & Ogles, B. M. (2004). The Efficacy and Effectiveness of Psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's Handbook of Psychotherapy and Behaviour Change* (5th ed.). New York: Wiley.
- Mellor-Clark, J., Connell, J., Barkham, M. & Cummins, P. (2001). Counselling Outcomes in Primary Health Care: a CORE System Data Profile. *European Journal of Psychotherapy & Counselling*, 4 (1), 65–86.
- Nathan, P. E., Stuart, S. P. & Dolan, S. L. (2000). Research on Psychotherapy Efficacy and Effectiveness: Between Scylla and Charybdis? *Psychological Bulletin*, 126 (6), 964–981.
- NHS (2011) <http://www.iapt.nhs.uk/about-iapt/> accessed 22/7/2011.
- Rowland, N., Bower, P., Mellor-Clark, J., Heywood, P. & Hardy, R. (2000). Counselling in Primary Care: a Systematic Review of the Research Evidence. *British Journal of Guidance & Counselling*, 28 (2), 215–231.
- Smith, M. L. & Glass, G. V. (1977). Meta-Analysis of Psychotherapy Outcome Studies. *American Psychologist*, 32 (9), 752–760.
- Spitzer, R. L., Kroenke, R., Williams, J. B. & Lowe, B. (2006). A Brief Measure for Assessing Generalized Anxiety Disorder: the GAD-7. *Archives of Internal Medicine*, 166, 1092–1097.
- Stiles, W. B., Barkham, M., Twigg, E., Mellor-Clark, J. & Cooper, M. (2006). Effectiveness of Cognitive - Behavioural, Person-Centred and Psychodynamic Therapies as Practised in the UK National Health Service Settings. *Psychological Medicine*, 36, 555–566.
- Stiles, W. B., Barkham, M. Mellor-Clark, J., & Connell, J. (2008). Effectiveness of Cognitive-Behavioural, Person-Centred, and Psychodynamic Therapies in UK Primary-Care Routine Practice: Replication in a Larger Sample. *Psychological Medicine*, 38 (5), 677–688.
- Wampold, B. (2001). *The Great Psychotherapy Debate*. New Jersey: Lawrence Erlbaum Associates.

## APPENDIX A: Adherence Questionnaires:

Both questionnaires carried the following instructions:

*This form should be completed every four weeks for each client using a segment of tape from the client's sessions.*

*Please complete this form by placing a cross in the box beside the number that you feel best represents the extent to which the practitioner adheres to each aspect of the model for Transactional Analysis. For this scale, the number 1 represents 'no adherence' and the number 5 represents 'Fully adhered', for example: -*

*Capacity to initiate, develop and maintain an effective therapeutic alliance, based on principles of respect and equality (I'm OK You're OK)*

No adherence	Little adherence	Partly adhered	Mostly adhered	Fully adhered
1	2	3	4	5

### Transactional Analysis Adherence Questionnaire

#### Section 1– Working Alliance, diagnosis, contracting and treatment planning and interventions

1. Capacity to initiate, develop and maintain an effective therapeutic alliance, based on principles of respect and equality (I'm OK You're OK)
2. Capacity to make an appropriate clinical assessment using Transactional Analysis theory
3. Capacity to make an appropriate clinical assessment using Core, DSM and/or where appropriate other diagnostic systems
4. Capacity to develop therapeutic contracts which take account of the needs and context of the client
5. Capacity to devise treatment plans based on Transactional Analysis models of treatment planning
6. Capacity to use the Therapeutic Operations to effect decontamination
7. Capacity to engage effectively with the co-transferential relationship as a means of effecting deconfusion

#### Section 2 – Theoretical Understanding and Reflection

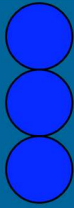
8. Capacity to describe and reflect upon clinical interventions using Transactional Analysis e.g. Game's analysis, Script analysis, TA proper, Ego state exploration, etc
9. Capacity to reflect on the effectiveness of interventions given the stage of treatment and the clinical content
10. Capacity to explain interventions using a variety of TA approaches
11. Can explain chosen interventions in response to an evolving sense of a personal style
12. Awareness of and capacity to reflect on own counter-transferential process and its meaning for the work

#### Section 3 – Working ethically, professionally and safely

13. Ability to work with issues of difference and pay attention to psychosocial, cultural and contextual factors as appropriate
14. Ability to engage in and effect risk assessments and address issues of safety as appropriate
15. Pays attention to ethical and professional issues and demonstrates ability to work with these
16. Uses appropriate professional support for the ongoing development of thinking and practice

## Integrative Counselling Psychology Adherence Questionnaire

1. Ability to conduct appropriate psychological assessments based on good inter-personal capabilities and a broad understanding of mental health
2. Ability to initiate, develop and maintain an effective therapeutic alliance
3. Demonstration of an understanding of psychopathology and diagnostic systems
4. A clear conceptualisation of treatment planning, goals and relevant change processes
5. Demonstration of the ability to contract with the client on therapeutic goals, activities and outcomes
6. Understanding of relationship dynamics at multiple levels of exchange
7. Demonstration of knowledge of theories of development throughout the lifespan
9. The capacity to integrate in a coherent way theories and competencies from more than one tradition in the psychological therapies
10. A capacity to attend to explicit and implicit communications and an ability to work with these
11. Sensitivity to attunement/misattunements
12. The ability to work with an understanding of the self in its multiple facets
13. Awareness of and capacity to reflect on own counter-transferential process and its meaning for the work
14. Understanding of the co-created nature of the therapeutic exchange
15. Effective and creative use of the self of the therapist
16. Ability to respond to complex demands as required
17. A capacity to attend to psychosocial, cultural and contextual factors as appropriate
18. Ability to work with issues of difference
19. A capacity to work towards self understanding in the client and an increase of awareness in the client of options for change
20. The ability to reflect on the appropriateness of interventions in line with stage of treatment, clinical content and client feedback
21. Engagement in risk assessment and attention to safety as appropriate
22. The capacity to attend to ethical and professional issues and the ability to work with these the use of appropriate professional support for the ongoing development of thinking and practice
23. The ability to monitor and evaluate therapeutic practice
24. A capacity to manage endings in the therapeutic process



## The Impact on Self Perception of Ego States of a Transactional Analysis Introductory Training Course (TA 101)

© 2011 Traian Bossenmayer

### Abstract

*The research examines the effects of transactional analysis (TA) 101 training upon self perceptions of ego-state dynamics, using the model of ego states incorporated into the Adjective Check List (Gough & Heilbrun, 1980). Subjects completed the questionnaires at the beginning and end of the training and one month later. The only statistically significant change was that Critical Parent decreased after the training and was still lowered one month later, although not as much. It was also found that gender was significant, but age was not.*

### Key words

transactional analysis (TA), TA 101, ego states, Critical Parent, Adjective Check List.

### Introduction

Many of the problems that occur in organisations are the direct result of people failing to communicate. Faulty communication causes a series of problems and it can lead to confusion and cause a good plan to fail (Pearson, 1983). Employees who are trained in transactional analysis theory might develop their skills in analysing transactional patterns and could be able to understand, predict and help improve dysfunctional, unproductive, uncooperative interactions between them-selves and their colleagues. In the TA literature it is assumed that TA training can help them communicate clearly and effectively at the three levels of the Parent (values) the Adult (rationality) and the Child (emotions, creativity) (Steiner, 1994).

The goal of this research was to test whether people who participated in a TA 101 training session (the introductory course in transactional analysis) had any changes in their pattern of ego-state behaviour. The research focused on the differences that might occur for each separate ego state at three different points in time. The positive effects of group therapy on ego-state change and ego-state perception was confirmed by Boholst's (2003) research on a group of 28 university students using the Adjective Check List questionnaire, the same method used in this research. The TA 101 course was chosen because of its standardised content and requirement that it is run

only by internationally-endorsed trainers, under the overall control of the International Transactional Analysis Association (ITAA) and the European Association for Transactional Analysis (EATA).

The majority of the participants of the TA 101 course are new to TA concepts, so the effects of the TA training can be observed in a more effective way. For this specific research a control variable was used to evaluate the differences between the TA 101 participants who already had some TA knowledge (from previous diverse trainings) and those who were encountering TA for the first time.

### Ego states: description and relevance for TA

Ego states represent one of the building blocks of TA theory. All transactional analysts work with ego states, which cover important personality features and are considered to be essential characteristics of TA therapy (Dusay, 1986).

In the early three ego-state models, the Parent is a language of values, the Adult is a language of logic and rationality, and the Child is a language of emotions. Creating an effective communication depends on the availability of all three intact ego states (Steiner, 1994). However, there are various models of ego states in use (Erskine & Trautmann, 1981, 1988; Van Beekum, 1996; Hargaden & Sills 2002; Hay, 2009) and it is recognised that this study is focused on behavioural diagnosis of ego states only.

The questionnaire used (Williams & Williams, 1980) is divided into five parts corresponding to the functional five ego-state model. Each ego state is regarded as a system of communication with its own distinct language and function (Steiner, 2011).

### Research question

The research set out to test the effects of the TA 101 training on the ego states of the participants by measuring the ego states at three different points in time: right before the training (T0), right after the training session ended on the second day (T1), and one month after the training (T2).

Based on this overall question, three hypotheses were derived:

Hypothesis 1. *There will be a change of the dominant ego states from T0 to T1*

Hypothesis 2. *There will be a change of the dominant ego states from T1 to T2*

Hypothesis 3. *There will be a change of the dominant ego states from T0 to T2*

These overall hypotheses converted into five sets to match the five ego states that were being measured in the questionnaire used.

Three control variables were used: the age of the respondents, the gender, and whether they had any past training in TA. It was noted that the way that ego states manifest is not related to age. This means that a 60-year-old person can act as a Free Child and a 12-year-old child can act as a Critical Parent.

### Research design

The research design was a pre-and post-intervention study with a follow-up measurement to check stability of change. The ego states of the subjects were measured before the training (T0), at the end of the training (T1) and one month after the training had ended (T2). The data was collected using a questionnaire and the participants were chosen from three different TA training sessions, all of them taking place in The Netherlands. Three control variables were used to get a better insight into the factors influencing these changes. The hypotheses were analysed using the student t test. To analyse the influences of the control variables on the differences between the means of each ego state, a multiple linear regression was used. Prior to this a bivariate analysis was used to check for correlations between the control variables and the differences in the ego state between the different points in time.

### Data collection

The assessed variable was the distribution of ego states for each participant at each of the time-points, measured using the TA sub-scales developed by Williams & Williams (1980) from the Adjective Check List (Gough & Heilbrun, 1980). Each participant received a list of adjectives which she/he scored as being characteristic for her/him; each ego state was derived by combining some of these adjectives; for instance, the Free Child ego state is a combination of 13 adjectives (e.g. adventurous, imaginative). The use of the Adjective Check List (ADL) scale was considered appropriate because adjectives were also commonly used by Berne (1967) to describe ego states. Also, this scale is academically validated.

The first two series of questionnaires, corresponding to T0 and T1, were collected directly from the trainers at the end of the sessions or were sent by mail. The third series, corresponding to T2, were collected by

sending individual emails to each training participant. The rate of response in T2 was lower compared to T0 and T1.

### Data analysis

Given the fact that the same subjects were analysed three times and that the interest was to see if there were any differences from one time point to another in the ego states, a multiple linear regression was used. This offered the possibility of analysing the interaction of the control variables (age, gender, prior TA training) on each ego state. Before this technique was applied, the average of each scale for each ego state per individual was calculated (for T0, T1, T2). Then the difference between the means (T1-T0, T2-T0, T2-T1) was used as a dependent variable in the multiple regression, with Cronbach's alpha of 0.6 applied as a criterion.

### Sample strategy

A random sample is desirable in any research in order to be able to generalise the results to the whole population from which the sample has been drawn. However, this was not possible in this study. The sample consisted of 38 people who were participating in the TA 101 trainings in The Netherlands in recent months. It had been hoped to have a sample of 100 people but this was not possible due to the lack of availability of TA 101 courses running during the limited time available for the research. The criteria for selection were the accessibility and the openness for medium-term investigation of the participants.

## Research quality indicators

### Reliability

Statistical reliability measures were used to ensure the reliability of the scales within the questionnaire (Cronbach's alpha 0.6). The reliability of the research is increased by the use of a standardised questionnaire and also by the standardised TA 101 training context and procedures which ensure a broadly similar learning experience and content for the participants.

### Internal validity

One way in which the quality of the research can be verified is by examining the psychometrical proprieties of the questionnaire. In this sense the distinctiveness sections measuring the five different ego states can be checked with the use of factorial analysis. Also, the questionnaire was academically validated through the study of Williams (1980) by using an expert panel of 15 transactional analysts. It was confirmed that each section was referring to a separate ego state and that the results of the questionnaire could be represented by the egogram developed by Dusay (1972).

Williams also mentions that the questionnaire was designed in order to offer an alternative for evaluating the strength of the ego states of participants who were

not familiar with TA concepts. Up to that point all the instruments of measuring the ego states were based on TA knowledge of the respondents. Taking into account that the majority of the TA 101 participants are new to the field of transactional analysis, the Adjective Check List questionnaire was the best suited method. The control variable of prior TA training was introduced in the study for the same reason. From the total number of 38 participants, only five had prior TA knowledge.

## Results

The five hypotheses for each time frame focus on the changes of each ego state from T0 (before the training), T1 (right after the course has ended) and T2 (one month after the training has ended). The results shown here are for Critical Parent only; the same processes were completed for each ego state but are not shown as they were not statistically significant. The data are available from the author for future researchers.

### Descriptive statistics

Table 1 includes the information about the TA 101 training regarding the number of participants, gender, age, and prior TA training. There were 38 participants

initially, ranging in age from 20 to 58 years, with an average age of 40 years. Only five of them had previous knowledge of TA concepts from other training sessions (not TA 101).

The data was collected from three different TA 101 training courses. For T0 (before the training) the total number of 38 participants (30 female, 8 male) filled in the questionnaires. For T1 (right after the training) there were 31 (24 female, 7 male) answered questionnaires because one trainer did not hand in their questionnaires. For T2 (one month after each training session) the participants were contacted by email. Only ten participants (3 female, 7 male) answered the emails and filled in the questionnaires.

Therefore the data for analysing the ego-state changes between T0 (before the training) and T1 (right after the session) come from 31 participants. The data for analysing the changes between T0 (before the training) and T2 (one month after the training) come from ten participants (maximum number of questionnaires for T2). The data for analysing the changes between T1 (right after the session) and T2 (one month after the training) come also from ten participants.

*Table 1: Descriptive statistics*

	Number	Minimum	Maximum	Mean	Std. deviation
gender	38	1,00	2,00	1,2105	,41315
age	38	20,00	58,00	40,1316	10,31968
TA training	38	1,00	2,00	1,8684	,34257
CP T0	38	1,15	3,77	2,4798	,55569
NP T0	38	2,54	4,77	3,7773	,47141
AD T0	38	2,77	4,46	3,5162	,41401
AC T0	38	1,46	3,38	2,4717	,52361
FC T0	38	1,92	4,42	3,4035	,56404
CP T1	31	1,38	3,92	2,4194	,62044
NP T1	31	2,69	4,77	3,8362	,47325
AD T1	31	2,77	4,38	3,5112	,45184
AC T1	31	1,38	3,23	2,3573	,48996
FC T1	31	2,33	4,50	3,4892	,42259
CP T2	10	1,08	3,15	2,0846	,63479
NP T2	10	3,00	4,92	3,7846	,55919
AD T2	10	2,54	4,08	3,4846	,46161
FC T2	10	3,00	4,17	3,4417	,35366
AC T2	10	1,23	3,08	2,1692	,64133

*Note.* CP = Critical Parent, NP = Nurturing Parent, AD = Adult, AC = Adapted Child, FC = Free Child



The scales are valid, the alpha Cronbach for all the five types of ego states are above 0.6. The data for the three time points are shown in Table 2.

Table 2: Alpha Cronbach for T0, T1, T2

	T0	T1	T2
<b>Critical Parent</b>	0.831	0.904	0.930
<b>Nurturing Parent</b>	0.851	0.894	0.935
<b>Adapted Child</b>	0.711	0.823	0.860
<b>Free Child</b>	0.836	0.756	0.654
<b>Adult</b>	0.824	0.831	0.922

### Analysis of the hypotheses

Before analysing each hypothesis using the student t-test, a bivariate analysis was used to compare the control variables (gender, age, prior TA training) with the changes in each ego state, to see if the variations in the values of the control variables are systematically associated with the variations in the ego states and to get a better insight on the interaction effect between them.

Table 3 includes the correlations between age and prior TA training and changes for each ego state at the three different points in time (T0, T1 and T2). The calculation for gender could be done at T0 only.

As we can see, for the Critical Parent the control variable gender was significantly correlated with the changes in the ego state. Therefore, in the Critical Parent results section, below, a separate t test analysis was run to see if the scores of the Critical Parent changed in a different way among males and females and whether these changes were significant or not. The results of the student t-test are shown in Table 4

For the Adapted Child the control variable age was significantly correlated with the changes in the ego state but the number of respondents in this case was low (10) so it was not feasible to split the sample into different age groups and compare them.

Table 3: Control variables – changes in the Critical Parent (CP)

		CP T0	CP T1	CP T2
<b>gender</b>	Pearson Correlation	.482**	.a	.a
	Sig. (2-tailed)	.006	.000	.000
	Number	31	10	10
<b>age</b>	Pearson Correlation	.257	.017	-.093
	Sig. (2-tailed)	.163	.963	.799
	Number	31	10	10
<b>TA training</b>	Pearson Correlation	.156	.041	-.062
	Sig. (2-tailed)	.403	.911	.864
	Number	31	10	10

Note. \*\* Correlation is significant at the 0.01 level (2-tailed);  
 .a cannot be computed because at least one of the variables is constant.

### Changes in the Critical Parent

**Hypothesis 1:** There will be a drop in Critical Parent (CP) ego-state at T1 (end of the training) compared to T0 (beginning of the training).

Hypothesis 0: CP T0 = CP T1

Hypothesis 1: CP T0 > CP T1

The  $t = 2.473$  from the table is higher than the standard value of  $t$  when Alpha is .05 and also that  $p$  of .019 is lower than .05, the null hypothesis is rejected. Therefore the analysis confirms with a probability of 95% that there is a drop of the Critical Parent ego state from T0 to T1.

The decrease of the ego state from T0 to T1 is also shown in Figure 1, the 'Critical Parent' graph and the difference of means is 0.17.

A multiple linear regression was completed to analyse the effects of the control variables on the decrease of the Critical Parent from T0 to T1. The F value from the ANOVA table is 3.468. The value for  $F_{0.05; 3.27} = 2.96$  (NIST/SEMATECH e-Handbook of Statistical Methods) is lower than 3.468. This means that the

Table 4: Paired samples test

		Mean	Std. deviation	Std. error	Lower	Upper	t	d	Sig. (2-tailed)
<b>Pair 1</b>	CP T2 – CP T0	-.35385	.39921	.12624	-.63942	-.06827	-2.803	9	.021
<b>Pair 2</b>	CP T1 – CP T2	.20000	.39089	.12361	-.07963	.47963	1.618	9	.140
<b>Pair 3</b>	CP T0 – CP T1	.14640	.32958	.05919	.02551	.26729	2.473	3 0	.019

95% Confidence interval of the difference

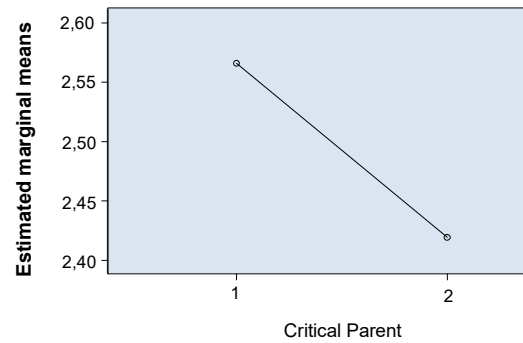
hypothesis of the multiple regression is confirmed and that the model is statistically valid.

For the change of the Critical Parent from T0 to T1 the effect of gender and constant is statistically significant and the effect of age and prior TA training is insignificant. The interaction between the constant, representing other variables that weren't controlled for, explains for .976 of the Critical Parent change.

Gender has a direct effect on the change and explains .355 of the change. Females had an average decrease of the mean of 0.15 and the males had an average increase of the mean of 0.21. The changes on both genders were significant.

Based on the fact that the number of female respondents decreased more from T0 to T1 than the number of males, the females had a stronger influence on the overall Critical Parent decrease. A representation of these results is shown in the following tables.

Figure 1: Estimated marginal means of MEASURE 1: Critical Parent T0-T1



Note. Covariates appearing in the model are evaluated at the following values: gender = 1.2258, age = 41.0968, TA training = 1.5387

Table 5a: ANOVA<sup>b</sup>

Model	Sum of squares	df	Mean square	F	Sig.
Regression	.906	3	.302	3.468	.030 <sup>a</sup>
1 Residual	2.352	27	.087		
Total	3.259	30			

Note. <sup>a</sup> Predictors: (Constant), TA training, gender, age; <sup>b</sup> Dependent Variable: CP1 0

Table 5b: Coefficients <sup>a</sup>

Model	Unstandardised coefficients		Standardised coefficients	t	Sig.
	B	Std. error	Beta		
(Constant)	-.976	.342		-2.858	.008
1 gender	.355	.128	.458	2.781	.010
age	.006	.006	.171	.979	.336
TA training	.075	.153	.085	.490	.628

Note. <sup>a</sup> Dependent Variable: CP1 0

Table 6a: One-sample statistics <sup>a</sup> – female

	Number	Mean	Std. deviation	Std. error mean
CP T0	30	2.4333	.57261	.10454
CP T1	24	2.2885	.61965	.12649

Table 6b: One-sample test <sup>a</sup> – female

Test Value = 0						
					95% Confidence interval of the difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
CP T0	23.276	29	.000	2.43333	2.2195	2.6472
CP T1	18.093	23	.000	2.28846	2.0268	2.5501

Table 6c: One-sample statistics <sup>a</sup> – male

	Number	Mean	Std. deviation	Std. error mean
CP T0	8	2.6538	.47950	.16953
CP T1	7	2.8681	.38644	.14606

Table 6d: One-sample test <sup>a</sup> – male

Test Value = 0						
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence interval of the difference	
					Lower	Upper
CP T0	15.654	7	.000	2.65385	2.2530	3.0547
CP T1	19.636	6	.000	2.86813	2.5107	3.2255

**Hypothesis 2:** There will be a drop in Critical Parent (CP) ego-state at T2 (one month after the training) compared to T0 (beginning of the training).

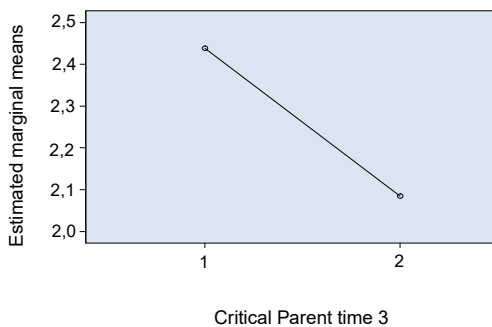
Hypothesis 0: CP T0 = CP T2

Hypothesis 1: CP T0 > CP T2

The paired sample test showed a  $t = -2.803$  for T0–T2, which is lower than the standard value of  $-t$  for Alpha = .05 and also because  $p$  of .021 is lower than .05, the null hypothesis is rejected. Therefore the analysis confirms with a probability of 95% that there is a drop of the Critical Parent ego state from T0 to T2.

The decrease of the ego state from T0 to T2 is shown in Figure 2, and the difference of means is 0.03.

Figure 2: Estimated marginal means of MEASURE 1: Critical Parent T0–T2



**Hypothesis 3:** There will not be any increase in Critical Parent (CP) ego-state at T2 (one month after the training) compared to T1 (end of the training).

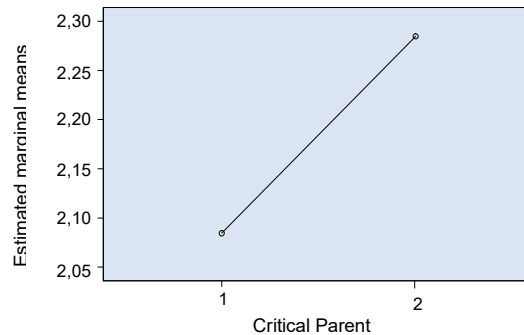
Hypothesis 0: CP T1 = CP T2

Hypothesis 1: CP T1 < CP T2

Considering the fact that  $t = 1.618$  is not higher than the value of  $t$  for Alpha = .05 and also that  $p$  at .140 is higher than .05, there is not enough evidence to reject the null hypothesis, so it is accepted. Therefore the hypothesis that there would not be an increase in Critical Parent from T2 to T1 cannot be statistically confirmed.

Table 3 shows there was an increase of Critical Parent from T1 to T2 of 0.22.

Figure 3: Estimated marginal means of MEASURE 1; Critical Parent T1–T2



## Discussion and limitations

### Discussion

If we take into account Steiner's (2003) argument that in a democratic, cooperative society it is recommended to limit the Critical Parent's control of human affairs, then the TA 101 can be confirmed as a valid instrument that can be used to obtain such a change. The Critical Parent ego state had a statistically confirmed drop after the training session (T1) and this drop was still maintained one month after the session had ended. The research could not confirm the fact that there was no increase of the Critical Parent during the month that passed before the final questionnaires were filled in (T1–T2). If we consider the graph of the changes between T1–T2 we can actually see an increase of the ego state from T1 to T2.

The multiple regression results show that gender had a direct effect on changing the Critical Parent. There was a different pattern of change for each gender, with females showing decreased use of Critical Parent whilst males showing an increase. These differences in ego-state change might be explained by the way participants conformed to gender-specific messages

and norms, which begin during childhood. Maltz and Borke's (1982) research showed that the games children play contribute to socialising children into masculine and feminine cultures. The Critical Parent ego state uses a very strong and strict communication style, which from a social point of view is perceived to be more masculine, while the Nurturing Parent uses a more emotional way of expression that is often regarded as more feminine. For example, according to Tannen (1990) women tend to express agreement and support more often than men, a characteristic that fits with the Nurturing Parent.

Williams (1980) follows the assumption of many TA theorists and states that there is a positive relationship between Critical Parent and Adapted Child. This means that as energy in the Critical Parent decreases so does the psychic energy in the Adapted Child ego state. This research has confirmed a decrease of the Critical Parent after the TA 101 training, but the decrease of the Adapted Child was not statistically significant. If we consider the graphs as evidence of the Adapted Child ego state drop we could argue that there is a decrease between T0-T1 for both Critical Parent and Adapted Child. However this positive relation is not confirmed for T0-T2 because the Adapted Child increases, while the Critical Parent ego state drops.

In this study only three control variables were included. However, the multiple regression used for Critical Parent from T0 to T1 showed that the changes in the ego state were influenced also by other factors that were not controlled for (constant in the coefficients table).

### Implications

The main implication of this research can be considered the empirical study on the effects of TA 101 course on the ego states. During the literature review process for this research no other studies could be found that focused on the effectiveness of the TA 101 course. There are a series of empirical studies (Rosenthal, 2000) regarding a similar process which focuses on transactional analysis therapy, which showed positive results.

### Limitations

It should be noted that this study has several limitations. The most significant one is the number of participants of the TA 101 training that filled in the questionnaires. The initial number was 38; after the training session 31 people filled in the questionnaire; one month later only ten participants responded.

It may also be that similar changes in ego states might result from any course of similar length that focused on teaching psychology to participants.

Regression analysis is normally performed on observed variables; using it here on a created variable of the

differences between the means requires that caution must be exercised in applying the results. The different group sizes for the T-test analysis relating to the gender variable also means that caution is called for.

Time was also a limitation of this research because the results had to be collected within three months, from March to May. Since TA 101 trainings are not very common in The Netherlands, finding participants willing to support this research in such a short time was challenging.

In addition, although the TA 101 has a standard syllabus and is run only by qualified trainers, there may well have been differences such as the emphasis placed on the various TA concepts, the models used by the trainers to explain ego states, the permission-giving (Crossman 1966) qualities of the trainers, the relationships the trainers formed with participants, and so on. The composition of the training groups may also have introduced further variables.

Another limitation concerns the cultural background of the participants. This research was carried out in the Netherlands, so certain characteristics of Dutch culture might have influenced the results.

### Direction for further research

The fact that this research could not prove that there was a significant change in self-perceptions of four of the five ego states does not necessarily mean that the TA 101 training is ineffective for this purpose. Taking into account the limitations of this study, a longer time for the research might offer the chance to include a larger sample of TA 101 participants.

Another direction for this research might be in testing the effects of TA 101 training directly in an organisational context by including members of the same teams in the course. Knowing each other and having already a relative bond at the beginning of the training, the participants could have better results in applying TA concepts and consequently changing their ego states in the desired direction.

In order to get a better perspective on the interaction between the Critical Parent and the Adapted Child, or between the Nurturing Parent and the Adult, an interesting direction for further research is to study also the interaction between all five ego states. Since these ego states are interconnected, a decrease in one ego state should mean an increase of at least another ego state, but exactly how these changes occur between ego states is still under discussion. Williams (1980) has tested and confirmed some of the interactions initially assumed by Berne, but his findings concerned transactional analysis therapy, not a TA 101 training process.

Finally, future research needs to take into account the plethora of ego-state models that now exist within TA

literature; the Adjective Check List was developed several years ago against a specific ego-state model that may now be out-dated and which may not have reflected the way ego states are currently taught on TA 101 programmes.

**Traian Bossenmayer**, MSc Organization Development, HR consultant and trainer, is in TA training as a transactional analyst (Organisational) and can be contacted on [traian.bossenmayer@gmail.com](mailto:traian.bossenmayer@gmail.com)

## References

- Berne Eric (1967). *Transactional Analysis in Psychotherapy*. New York: Atherton Press.
- Crossman, P. (1966). Permission and Protection *Transactional Analysis Bulletin*, 5 (19), 152–154.
- Dusay, J. M. (1972). Egograms and the constancy hypothesis. *Transactional Analysis Journal*, 37–41.
- Dusay, J. M. (1986). Transactional analysis. In I. L. Kutash & A. Wolf (Eds.), *Psychotherapist's casebook*, 413–423. San Francisco: Jossey-Bass.
- Erskine, R. & Trautmann (1981). Ego State Analysis: A Comparative View. *Transactional Analysis Journal*, 11 (2), 178–185.
- Erskine, R. & Trautmann (1988). Ego Structure, Intrapsychic Function and Defense Mechanisms: A Commentary on Eric Berne's Original Theoretical Concepts. *Transactional Analysis Journal*, 18 (1), 15–19.
- Gough & Heilbrun (1980). *The Adjective Check List Manual*. Palo Alto: Consulting Psychologist Press.
- Hargaden, H. & Sills, C. (2002). *Transactional Analysis A Relational Perspective*. Hove: Brunner-Routledge.
- Hay, Julie (2009). *Transactional analysis for trainers*. Hertford: Sherwood Publishing.
- Maltz, D. & Borker, R. (1982). A cultural approach to male-female miscommunication. In J. Gumperz (Ed.), *Language and social identity*, 196–216. Cambridge, UK: Cambridge University Press.
- NIST/SEMATECH e-Handbook of Statistical Methods*, <http://www.itl.nist.gov/div898/handbook/eda/section3/eda3673.html> Accessed on 31/12/2010.
- Pearson, J. (1983). *Interpersonal Communication*. Glenview, Illinois: Scott, Foreman and Company.
- Rosenthal, W. S. (2000). Relationship of Change in Ego-State to Outcome of Stuttering Therapy: Preliminary Findings. Hayward: California State University.
- Steiner C. (2003). *Emotional Literacy; Intelligence with a Heart*. Fawnskin, CA: Personhood Press.
- Tannen, Deborah (1990, June 24). Sex, Lies and Conversation; Why Is It So Hard for Men and Women to Talk to Each Other? *The Washington Post*.
- Van Beekum, Servaas (1996). The graphics of Ego States. Advanced Working Conference, Amsterdam, Netherlands.
- Williams, K. B. & Williams, J. E. (1980). The assessment of transactional analysis ego states via the Adjective Check List. *Journal of Personality Assessment*, 2, 120–129.