MODUMBAD
- en kilde til liv

UNIVERSITETET I OSLO

Forskningsinstituttet Modum Bad 2010

MOVING THE MOUNTAIN:
Making psychotherapy process research time and cost effective.

PÅL ULVENES, LENE BERGGRAF, ASLE HOFFART AND LEIGH MCCULLOUGH

INTRODUCTION:

Understanding what makes therapy effective is an important next step and at the cutting edge of psychotherapy research (Kazdin, 2007, Annu. Rev. Clin. Psychol). However, there are enormous obstacles in conducting this kind of research. The study of process is tremendously complex as well as time consuming, labor intensive, and hugely costly. Therefore, process researchers often feel like they are trying to 'move a mountain.'

This project builds on a former study (Berggraf, 2010) that examined ways to streamline process research. This study controls for methodological limitations in the earlier study; controlling for practice effects, type of therapy, type of patients and increasing the sample size.

METHOD:

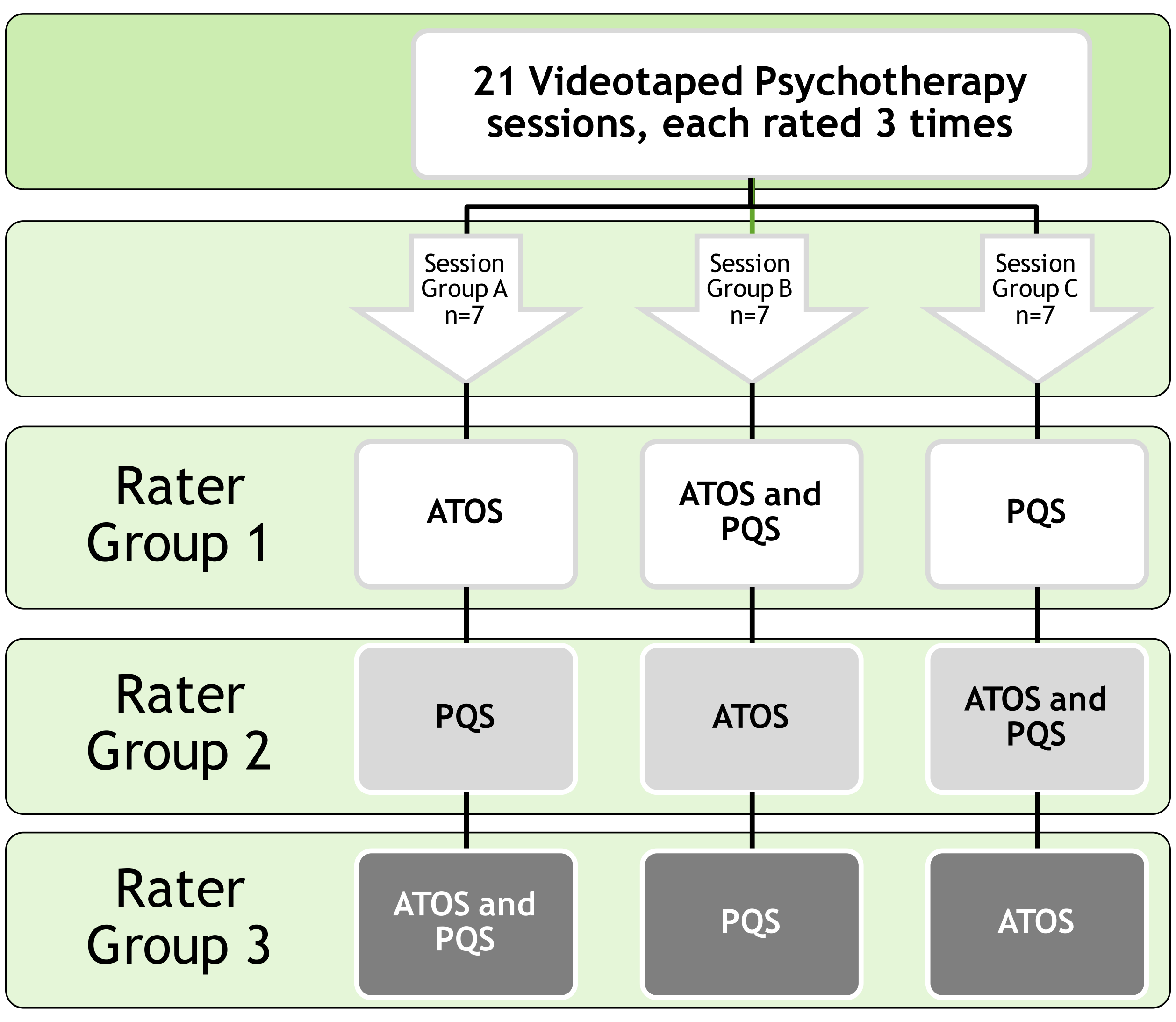
This project investigates whether raters can simultaneously code more than one process measure without one measure confounding the other.

The ATOS, an in-depth measure of patient change, is scored on 5 subscales every 10 minutes of the session and on 2 subscales at the end. The PQS, a broader descriptive measure of therapist-patient interaction, is rated at the end of a session. Both instruments are well operationalized and behaviorally grounded.

Thirteen reliable raters (ATOS, ICC> .7; PQS, r> .5) was randomly assigned to 3 rater groups and coded 21 psychotherapy sessions from the Svartberg et al. (2004, Am J Psychiatry) RCT. Two instruments were used; ATOS: Achievement of Therapeutic Objective Scale & PQS: Psychotherapy Process Q-sort.

Each Rater Group rated 21 sessions in three batches of 7 sessions each. The Rater Groups rated the batches in one of three conditions (ATOS alone, PQS alone, and ATOS and PQS together). The sessions were counterbalanced for order of presentation. Thus, each of the 21 sessions was rated three times, once in each of the conditions, and no Rater Group saw the same session twice (See research design in Figure 1).

Figure 1: Design of Study



21 Videotaped Psychotherapy sessions, each rated 3 times

Session Group A n=7 | Session Group B n=7 | Session Group C n=7

Rater Group 1: ATOS | ATOS and PQS | PQS

Rater Group 2: PQS | ATOS | ATOS and PQS

Rater Group 3: ATOS and PQS | PQS | ATOS

Generalizability Theory (GT):

GT improves upon Intraclass Correlation (ICC) with a more sophisticated analysis, and provides two new statistics:

1) GT Variance Components: GT is analogous to ANOVA because it partials out the variance attributable to different factors ('facets' in GT); e.g, raters, patients, etc. (Brennan, 1992, Am Coll Testing). In ICC, the contributions from these factors are combined into one statistic, so it is not possible to know what portion of the variance is due to each factor (raters, patients, etc). Thus, GT is a more precise analysis than ICC, just as ANOVA is more precise than a t-test.

2) G-Coefficient: One function of GT is a Generalizability (G) Coefficient that provides a test of reliability - like an ICC. However, the G-coefficient takes more factors into account. Drawing from the variance components (#1 above) it weights each factor ('facet') according to the magnitude of its contribution to measurement error. Thus the G-coefficient is a more correct measure of reliability (or generalizability) than ICC.

RESULTS:

For 6 of the 7 scales the variance components show no confounding of measurements when multiple measures are rated together by highly trained raters (see table 1). The G-coefficients revealed very little difference between conditions of measurement (see table 2). The only scale to have a G-coefficient lower than .97 was New Learning (.20). Upon inspection this was due to changes in measurement procedures. This has been corrected and is being recoded.

DISCUSSION:

This study adds further support that coding of multiple constructs and instruments can be conducted when raters are highly trained and when measures are highly grounded in observable behavior.

This study combined with the previous study indicate remarkably similar findings even though there were very different samples of patients, therapists, and conditions. This enhances the generalizability of the findings.

The "Mountain-Mover Team" at Modum Bad.

Rows 1 &2: *Pal Ulvenes, Linn Nygaard, Johan Aronsen, Ida Marie Holt, Andreas Espetvedt, Saher Sourouri, Kjetil Bremer, Natalia Gits, Ann Mari Bardum, Eva Jonassen, Linn Kolstad, Christine, Lingås, Kjersti Togstad, Ingrid Malin, Mari Berg, Liv Norum, Anna Melkeråen, Silje Friberg, Lene Berggraf, Aslak Hjeltnes, Tuva Øktedalen, Stuart Ablon, Julie Ackerman, Ray Levy, Leigh McCullough.*
Row 3: *Kristoffer Nordheim, Heidi Steen, Trude Alsos, Andreas Segrov, Mari Asmyhr.*
Row 4: *Øystein Verås, Øyvind Bjørkum, Even Halland, Yngvild Aamodt, Vidar Sandaunet, Kristina Tønnesen, Dag Wennesland.*
Not pictured: *Espen Folmo, Gunnar Gjermundsen, Morten Hegdal, Jeanette Sjaaeng, Tiril Østefjells, John Winter, Tyra Tambwe, Alexander Holmboe, Chris Kåring, Ragnhild Vagen, Charlotte Kirkhus, Nina Tangstrøm, Tuva Langjord.*

Table 1: Per Cent of variance by Factors ('facets') for the ATOS Subscales*

	Patient	Time	Condition	Segment	Raters
Insight	29.3%	0.0	0.0	0.0	19.6%
Motivat'n	35.9%	0.0	0.0	0.0	47.3%
Activat'g Affect	30.7%	0.0	0.0	0.0	17.8%
Inhibitory Affect	35.7%	0.0	0.0	1.0%	53.1%
New Learning	11.4%	0.0	0.0	**	88.6%
Sense of Self	72.7%	0.0	0.2%	**	10.7%
Sense of Others	61.8%	0.0	3.5%	**	19.1%

*Only main effects are reported.

** Scales only rated at end of session. Segment factor is not applicable.

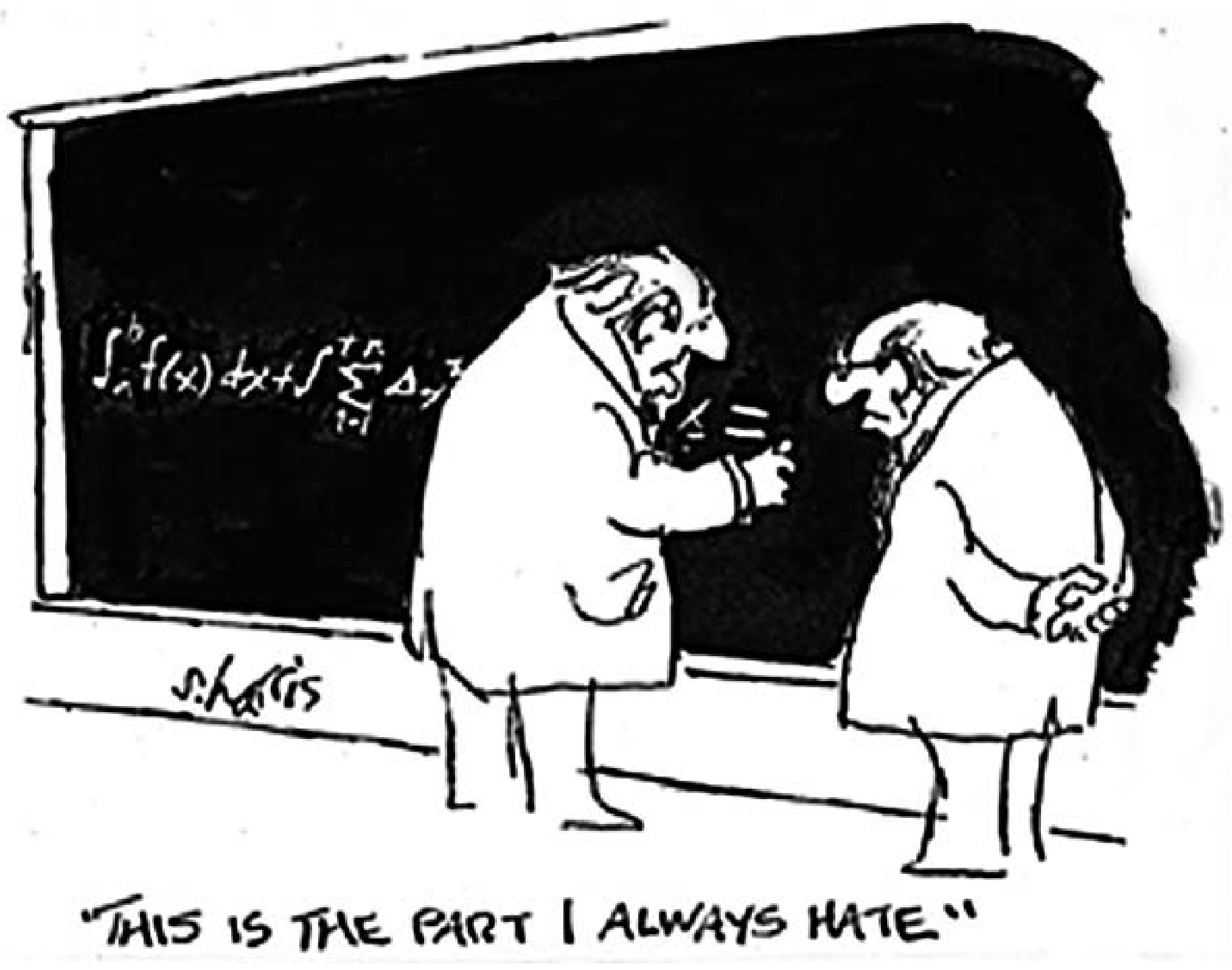

$\int_a^b f(x)\,dx + \int \sum_{i=1}^{\tau n} \Delta_y$
S.Harris
"THIS IS THE PART I ALWAYS HATE"

Table 2: Generalizability Coefficients

	Insight	Motivation	Activation	Inhibition
G	.99	.97	.99	.97

	New Learning	Sense of Self	Sense of Others	PQS
G	.20	.99	.97	.98

Limitations:

This study used only observer ratings. Future studies should include other process measures. Also, the sample is homogenous, and limited in size. Future research should adress this.

CONCLUSION:

While more studies are called for, we can begin to consider more streamlined or efficient ways of collecting process data. This will help us grow in understanding how psychotherapy best can help our patients.

PÅL ULVENES is a Psychologist and a Doctoral Candidate at Modum Bad Research Institute and the University of Oslo.