

# **Abschlussbericht des Modellversuches DINGI**

**KAPITEL 6      Auswertung**

**Dr. R. Thomann**

Königsfelden, im September 1998

## KAPITEL 6 AUSWERTUNG

(von Dr. med. Reinhold THOMANN, Psychiatrische Klinik Königsfelden)

### INHALTSVERZEICHNIS

#### 6.1 Allgemeines

##### 6.1.1 Methodik

#### 6.2. FPI-Teste

##### 6.2.1. Allgemeines

##### 6.2.2. Resultate

##### 6.2.3. Diskussion der FPI - Resultate

#### 6.3. GT-Test

##### 6.3.1. Allgemeines

##### 6.3.2. Resultate

##### 6.3.3. Diskussion der GT - Befunde

#### 6.4. U-Fragebogen

##### 6.4.1. Vorgehen

##### 6.4.2. Resultate

##### 6.4.3 Interpretation der U-Fragebogen-Resultate

#### 6.5 Token-Economy

##### 6.5.1 Vorgehen

##### 6.5.2. Resultate

##### 6.5.3 Diskussion der Token-Economy-Resultate:

#### 6.6. Katamnestische Erhebungen

#### 6.7. Zusammenfassung der Resultate, Diskussion und Ausblick

#### 6.8. Dank

#### 6.9. Beilagen

## PROJEKT DINGI

### 6. AUSWERTUNG

(von Dr. med. Reinhold THOMANN, Psychiatrische Klinik Königsfelden)

#### 6.1 Allgemeines

Normalerweise werden bei Forschungsarbeiten Hypothesen aufgestellt, die dann aufgrund der Resultate angenommen werden können, d.h. erfüllt werden oder aber verworfen werden müssen. Das heisst, dass im einfachsten Fall nur zwischen "erfüllt" oder "verworfen" entschieden werden muss. Hierzu sind schon vor der Durchführung entsprechender Versuche die "Messwerkzeuge respektive Messgrössen" zu definieren und ebenso das mathematische Verfahren, mit Hilfe dessen die Beurteilung der Hypothese erarbeitet wird.

Im vorliegenden Fall handelt es sich bei der Haupthypothese gemäss 3.1 ("je besser jemand sozialisiert ist, desto kleiner ist die Wahrscheinlichkeit einer zukünftigen Delinquenz bzw. desto grösser ist die Wahrscheinlichkeit einer Reduktion von Deliktschwere, Delikthäufigkeit sowie Rückfallgeschwindigkeit") um eine relativ komplexe Aussage, die nicht einfach mit ja oder nein beantwortet werden kann. Dies verrät schon die komparative Formulierungsart der Haupthypothese: "Je besser..., desto kleiner...". Zur Operationalisierung respektive "Messbarmachung" dieser Hypothese werden dann auch Bereiche "Sozialisation" und "Legalbewältigung" genannt. Unter 4.3.2 werden denn die eigentlichen "Messinstrumente" genannt, nicht erwähnt wird aber, welche Werte in welchem Mass und wie verändert sein müssen, dass die obgenannte Haupthypothese angenommen werden oder verworfen werden kann.

Es gibt aber auch viele Untersuchungs- und Auswertungsmöglichkeiten, bei denen eine strenge experimentelle Anordnung wie im vorgenannten Sinne nicht unbedingt notwendig ist. Dies gilt z.B. auch bei nachträglichen Auswertungen von Phänomenen, über die gewisse Aufzeichnungen existieren.

Ganz allgemein sind - wenn statistisch zulässig - Auswertungen von metrischen, d.h. in Form von Zahlenwerten ausgedrückten Parametern am empfindlichsten, d.h., mit ihnen kann man am ehesten eine Veränderung von zwei solchermassen erfassten Situationen statistisch berechnen. Dabei geht es jeweils ja darum, eine Aussage machen zu können, wie "signifikant", d.h. zu welchem Prozentsatz diese Veränderungen nicht auf Zufallsverteilung basieren.

Im FPI-Test (siehe 4.3.2.2), Giessen-Test (siehe 4.3.2.3) sowie U-Fragebogen sind die Bedingungen zur metrischen Vergleichsanalyse erfüllt.

### 6.1.1 Methodik

Aus den vielen erhobenen Fragebogen mussten die Daten zuerst einmal in Tabellenform "EDV-gerecht" aufbereitet werden. Dabei wurden die Daten gleichzeitig *anonymisiert*, d.h. für die weiteren Berechnungen wurden nur noch Versuchsnummern (für die Experimentalgruppe 201 - 218, für die Kontrollgruppe 101 - 114) zugeordnet. Zu diesem Zwecke standen Tabellenkalkulations-Programme (Lotus 1-2-3 Version 2.1, Excel Version 4 und später auch 5) zur Verfügung. Zur Datenauswertung konnten zum Teil schon in die Tabellenkalkulations-Programme eingebaute Funktionen verwendet werden, für spezielle Analysen wurde das Programm "STAT-GRAPHICS", Version 3 (STSC Inc., 1989) auf einer DOS-Plattform verwendet. Leider hat sich gezeigt, dass die Tabellenkalkulations-Programme in gewissen Konstellationen reproduzierbar fehlerhafte Resultate liefern, was dann zur Folge hatte, dass sämtliche Werte mittels letzteren Programmes nochmals durchgeprüft werden mussten. Da leider auch jetzt die Stichprobenzahl der Kontrollgruppen aus erhebungstechnischen Gründen nur auf 14 vermehrt werden konnte - gegenüber 18 in der Experimentalgruppe - sind direkte und evtl. auch paarweise Vergleiche zwischen den zwei Stichprobengruppen "KONTROLLE" und "EXPERIMENTAL" nicht möglich. Die zulässigen Tests beschränken sich auf numerische statistische paarweise Gegenüberstellung von erst und Zweitmessung *innerhalb* der beiden Stichproben und auf die Bewertung und Gegenüberstellung der dabei erhaltenen Signifikanzen und Trends sowohl bezüglich numerischer Grösse (skalar) als auch Abweichungsrichtung (vektoriell). Letztere sind, wenn sie zum Beispiel in den 2 Subsamples entgegengesetzt sind, auch bei zahlenmässig nur als Trend in Erscheinung tretenden Resultaten dann besonders wichtig. Deshalb wird auch weitgehend eine Diagrammdarstellung gewählt, die beide Aspekte zusammen leichter sichtbar werden lässt. Abb. 1c und 1d zeigen als Beispiel derart aufbereitete Datentabellen, hier dargestellt anhand der Statinenwerte für die FPI - Tests, weiteres siehe dort.

In den nachfolgenden Resultatbeschreibungen wird zur Textvereinfachung stets die Kontrollgruppe mit einem "K" und die Experimentalgruppe mit einem "E" dargestellt.

### 6.1.2 Bestimmung der möglichen Verteilung bei den Stichproben

Für FPI und GT darf eine Gaus'sche Normverteilung der Stichprobe angenommen werden. Die Rohdaten des FPI- und GTI-Testes (vorderhand nur der Experimentalgruppen) wurde einer entsprechenden exploratorischen Datenanalyse mittels STATGRAPHIC unterzogen. Dabei zeigte sich, dass die Normalverteilung gegenüber anderen geprüften Verteilungsmustern (geometrisch, exponential oder logarithmisch) am besten zu den Datenpunkten passten. Damit ist es auch zulässig, entsprechende Auswertungsmethoden wie nachfolgend genannt, anzuwenden.

## 6.2. FPI-Teste

### 6.2.1. Allgemeines

Nach graphischer Darstellung und Prüfung jeder einzelnen Datenreihe auf allfällige "Ausreisser" wurden die Rohwerte jedes Samples in die entsprechenden alters - und geschlechtsgematchten **STANINEN** umkodiert (siehe .Abb. 1c und 1d). Staninen bezeichnen die Ortsbereiche auf der Abszisse 3. Zeile innerhalb einer Normalverteilungskurve. Da erstens bei solchen wie auch ähnlichen Tests die Roh-Werte aufgrund verschiedener Fragen berechnet werden und bei der Testeichung dann bezüglich Validität und Verteilung entsprechend transformiert werden müssen, werden sie häufig in Form solcher "Standort-Staninen" angegeben. Wie Abb. 1e zeigt, gibt es insgesamt neun Stanine-Werte. Im Bereich der mittleren, d.h. 5. Stanine, befinden sich 20 % aller Resultate, im Bereich der 4. respektive 6. 17 und im Bereich der 2. respektive 7. 12 % etc. Da die Resultate hier auf einer linearen Skala aufgetragen werden (im Gegensatz z.B. zum GT-Test, siehe dort), ist diese Umkodierung nötig.

Erst aus diesen wurden Mittelwerte und Standardabweichungen berechnet. Die Graphische Darstellung beschränkt sich zu Gunsten besserer Uebersichtlichkeit auf Mittelwert abzügl. sowie zuzügl. halbe Standardabweichung, was ein "Profil" mit der **Breite einer Standardabweichung** ergibt. Durchgezogene Linien bezeichnen den Bereich der Erst-, gestrichelte diejenigen der Zweitmessung.

Mittels des Student-T-Testes wurde die Signifikanz der Veränderungen berechnet. Bei unserem Setting ist eine gepaarte Datenanalyse zulässig (stets gleiche Versuchsperson Messung 1 mit Messung 2 derselben Person). Da aufgrund der Hypothese aber nicht beschrieben worden ist, in welche Richtung

die allfällige Veränderung stattfinden müsste, konnte nur das symmetrische Verfahren verwendet werden, welches gegenüber asymmetrischen weniger schnell statistisch signifikante Werte ergibt.

Die p-Werte ("Signifikanzwerte") sind rechts für jedes "Item", nachfolgend **SS** (als Abkürzung für **Subskala**) genannt, als Zahlen angegeben; ein Pfeil markiert den Vektor der entsprechenden Veränderung. (Abb. 1a, 1b).

### 6.2.2. Resultate (siehe Abb. 1a und 1 b)

Bei **K** zeigt sich lediglich in **SS 7 "Beanspruchung"** eine allerdings nur *trendmässige* Veränderung in Richtung *"weniger beansprucht, gefordert und belastbar"* mit  $p = 0,12$

-----

Bei **E** zeigen sich hingegen bei **zwei SS** **signifikante Veränderungen**:

**SS 4 "Gehemmtheit"** in Richtung *"ungezwungener, selbstsicherer und kontaktbereiter"* mit  $p = 0,05$

und

**SS 6 "Aggressivität"** in Richtung *"aggressiveren, spontaneren aber auch reaktiveren - sich durchsetzenderen Verhaltens"* mit  $p = 0,04$ .

Trends zeigen

**SS 1 "Lebenszufriedenheit"** hin zu *"vermehrter Lebenszufriedenheit, guter Laune und Zuversichtlichkeit"*,  $p = 0,12$ .

**SS 2 "soziale Orientierung"** hin zu vermehrt sozialer Orientierung und mehr Eigenverantwortung mit  $p = 0,13$ .

**SS 3 "Leistungsorientierung"** in Richtung *"aktiver, schneller handelnd, ehrgeiziger, konkurrierender"* mit  $P = 0,12$

**SS 5 "Erregbarkeit"** in Richtung *"erregbarer, empfindlicher und unbeherrschter"*,  $p = 0,09$ .

-----

Während bei **K** lediglich **SS 7** einen Trend zeigte, sind bei **E** im Kontrast dazu *keine Veränderungen feststellbar* ( $p = 0,43$ ).

### 6.2.3. Diskussion der FPI - Resultate

Währenddem die nur trendmässige Veränderung bei Subgruppe 7 von **K** in Richtung verminderte Beanspruchung allenfalls mit einem möglichen

Prisonisierungseffekt vereinbar ist, im Rahmen dessen es durch vermehrte Habituation an ein regressives Regime gekommen ist, handeln könnte, zeigen sich doch bei der Experimentalgruppe deutliche Hinweise in Richtung verbesserter Sozialfähigkeit und mehr jetzt von innen heraus gesteuerte Verhaltensmuster. Die Veränderung auch in Richtung hin zu vermehrt Aggressivität ist natürlich nicht unbedingt wünschenswert, sie ist aber Ausdruck der zugrundeliegenden Persönlichkeitsstörung. Eine solche lässt sich in dieser Zeit wenig beeinflussen. Durch vermehrte emotionelle Ansprüche kommt es jetzt auch vermehrt zur Notwendigkeit, "Dampf ablassen" zu müssen. Während die anderen festgestellten Eigenschaften eigentlich gute Startbedingungen für die weitere Lebensbewältigung darstellen, ist mit diesem Item wahrscheinlich auch ein Anteil der tief zugrunde liegenden Persönlichkeits-Störung zutage getreten. Man könnte allerdings darin auch eine Möglichkeit sehen, dass diese Personen jetzt schneller ihren "Dampf ablassen" und vielleicht sozial zwar unangenehm auffallen, aber eventuell das Delinquieren lassen können. Hier wird erst die Beobachtung des katamnesticen Verlaufes Aufschluss geben.

Somit kann also festgestellt werden, dass bereits mit dem Werkzeug des FPI-Selbstbeurteilungstestes massive und eindeutige Unterschiede zwischen Kontroll- und Experimentalgruppe bestehen.

### **6.3. GT-Test**

#### **6.3.1. Allgemeines**

Das rechnerische Vorgehen entspricht weitgehend demjenigen des FPI-Testes, wobei aber die Rohwerte direkt zur weiteren Berechnung verwendet werden konnten und nicht erst noch in andere wie z.B. STANINEN umkodiert werden mussten. Dafür mussten die berechneten Werte auf das Original-Profilblatt von Hand eingezeichnet werden, weil in diesem Fall die Skalen nach "Eichung" erstens nicht linear und zweitens nicht spaltenweise, d.h. mit gleichen Werten untereinander, angeordnet sind. Dies entspricht in etwa dem Vorgang, der beim FPI mit der Umcodierung schon vorweggenommen worden ist und dafür zu linearen karthesischen Darstellungen führte.

### 6.3.2. Resultate (Abb. 2a und 2b)

*K* (ontrollgruppe) zeigt bei

**SS (HM-DE)** (Grundstimmung hypomanisch / depressiv) eine *signifikante Veränderung* in *Richtung hypomanisch* mit  $p = 0,04$ .

**SS (PO-IP)** (Soziale Potenz: Potent / Impotent) eine *stark trendmässige Aenderung* hin zu "*sozial weniger potent*" mit  $p = 0,07$ .

-----

In *E* (xperimentalgruppe) finden sich bei

**SS (UK-ZW)** (unkontrolliert / zwangshaft) eine *hochsignifikante Aenderung* in Richtung *vermehrt zwangshaft* mit  $p = 0,01$

**SS (PO-IP)** (siehe oben, soziale Potenz) eine *signifikante Aenderung* hin zu *grösserer sozialer Potenz* mit  $p = 0,05$ .

**SS (DO-GE)** (Gehemmtheit) *stark trendmässig* in Richtung *vermehrte Dominanz* mit  $p = 0,07$ .

**SS (HM-DE)** - Bedeutung siehe oben - angedeutet trendmässige Aenderung hin zu hypomanischerem Verhalten mit  $p = 0,12$ .

-----

### 6.3.3. Diskussion der GT - Befunde

Bei *K* zeigt sich bei **SS (HM-DE)** "Grundstimmung" eine signifikante Veränderung in Richtung hypomanisch ( $p = 0,04$ ) wogegen bei *E* nur *ein Trend* in dieselbe Richtung ( $p = 0,12$ ) zeigt.

Dieser Sachverhalt ist schwierig erklärbar. Evtl. kann man eine Hypothese aufstellen, wonach die Personen aus *K* durch den relativ starren, rigiden Tagesablauf und die verminderten emotionellen Belastungen sich eventuell weniger depressiv fühlen. Dabei darf nicht vergessen werden, dass es sich bei unserer Population um meistens schwer persönlichkeitsgestörte Individuen handelt. Entsprechend ist ein Vergleich mit normalpsychologischen Masstäben nicht sinnvoll. Die relativ emotionsarme Situation im Normalvollzug, wo keine

grösseren emotionalen Anstrengungen und auch Belastungen entstehen, könnte diesen Befund teilweise erklären.

Die trendmässige Änderung in Richtung "weniger sozial potent" bei **K** kann allenfalls Ausdruck eines Prisonisierungseffektes sein, insbesondere, wenn die **statistisch signifikante und gegenläufige Veränderung** bei **E** (in Richtung sozial potent) gegenübergestellt wird. Der Trend bei **E** in Richtung "dominanteres Verhalten" korreliert mit der im FPI Test gemachten Feststellung zu vermehrter Aggression, siehe auch dort.

Die Veränderung zu vermehrt zwanghaftem Verhalten (Trend,  $p = 0.01$ ) kann in diesem Rahmen auch positiv gewertet werden, indem eben Abmachungen "mehr zwanghaft" das heisst hier auch regelmässiger und zuverlässiger eingehalten werden können, somit besteht also ein Trend zu **vermehrter Selbstkontrolle**. Dies ist ja eine wesentliche Voraussetzung für ein erfolgreicherer, besser adaptiertes Sozialverhalten.

Während die soziale Potenz bei **K** abnimmt, steigt sie bei **E**.

Somit ergeben sich insgesamt Hinweise für deutliche Veränderungen bei der Experimentalgruppe in Richtung allgemein vermehrter sozialer Potenz. Das heisst vermehrt Selbstkontrolle und auch Bereitschaft zur Übernahme von Verantwortung.

## 6.4. U-Fragebogen

### 6.4.1. Vorgehen

Auch beim U-Fragebogen wurde dasselbe tabellarisch-mathematische Verfahren angewendet wie bei den vorhergehenden zwei Tests. Eine komprimierte grafische Darstellung ist hier gegenüber dem FPI und GT Test schwieriger, weil die einzelnen Subgruppen verschiedene Wertbereiche haben und somit kaum in einem einzigen Diagramm noch überschaubar dargestellt werden können. Entsprechend wurde hier eine EINZELDARSTELLUNG der einzelnen Subgruppen gewählt, woraus der Verlauf und die Streuung hervorgeht (Abb. 3).

### 6.4.2. Resultate

Auffallend sind die zum Teil stark abweichenden Anfangsstreuungen (Erstmessung - besonders Subskalen 1,2 und 5. Sie weisen auf **heterogene Ausgangsbedingungen** zwischen den beiden Unterstichproben hin. Alleine daraus muss der Schluss gezogen werden, dass beim Randomisierungsprozess ein sogenannter BIAS (= in etwa: Systematischer Fehler bei der Zuteilung zu den beiden Kategorien Experimental und Kontrollgruppe) ergeben haben muss. Beim Zurückverfolgen des experimentellen Ablaufes zeigte sich auch, dass es 1997/98 schwierig wurde, genügend Personen für die Kontrollgruppe ausscheiden zu können. Dabei wussten die beteiligten Probanden schon zum voraus, dass sie nie die Möglichkeit haben werden, in die Dingi-Gruppe eingeteilt zu werden.

Wie aus Abb. 3 hervorgeht, ergibt sich bei **K**

in **SS 4** (Nicht Neinsagen können) eine signifikante Aenderung hin zu kleineren Rohwerten entsprechend besserem NEINSAGEN mit  $p = 0,04$ .

in **SS 5** (Schuldgefühl) einen Trend zur Verminderung (höhere Rohwerte),  
 $p = 0,12$

Bei **E** zeigt sich nur bei

**SS 4** ein **Trend** mit  $p = 0,14$ , konkordant zu **K**.

Die Absolutwerte befinden sich bis auf wenige Ausnahmen stets im statistischen Normbereich. Lediglich bei **SS 3** (fordern können) zeigen sich bei fehlender rechnerischer Signifikanz Auffälligkeiten, indem bei **E** gegenüber **K** in der Zweitmessung die Streuung zugenommen hat. Eine vorsichtige Interpretation könnte darin bestehen, dass in der Dingi-Gruppe mehr individuelle Entwicklungen möglich waren, die zur grösseren Streuung bei M2 führten. Schon bei der orientierenden Voruntersuchung (siehe S. 21 ff) zeigten 18 von 19 Personen ein hochauffälliges, überschliessendes Sozialverhalten, so dass die aktuellen Befunde bezüglich Absolutwerten eher besser ausgefallen sind.

#### **6.4.3 Interpretation der U-Fragebogen-Resultate**

Wahrscheinlich sind die Veränderungen in **SS 4** Ausdruck des Prisonisierungseffektes, zudem zeigt sich aufgrund der Verteilung bei M1 (Erstmessung), dass die Randomisierung nicht optimal durchgeführt worden ist. Ebenso kann das verminderte Schuldgefühl (**SS 5**) bei **K** damit zusammenhängen, so hat man sich doch dort an die Umgebung einigermaßen angepasst und weiss auch, dass es vorderhand nichts zu verändern gibt. Warum bei **K** die Fähigkeit zum Neinsagen deutlich zugenommen hat und bei **E** nicht, muss offen bleiben.

## 6.5 Token-Economy

### 6.5.1 Vorgehen

Die Auswertung dieses "Werkzeuges" ist gegenüber den vorher geschilderten insofern anders, als es nur bei der Experimentalgruppe zum Einsatz gekommen ist, nicht aber bei der Kontrollgruppe und somit auch keine Gruppen-Vergleiche zulässt. Da die Token-Economy nicht nur ein Messinstrument (d.h. Erhebung von Parametern zu gewissen Zeitpunkten) darstellt, sondern auch ein das Verhalten beeinflussendes System darstellt, ist die Auswertung heikel, denn es gibt hier weder Signifikanzen noch Vergleichsmöglichkeiten.

Da nicht alle Versuchspersonen gleich lang im Dingi gewesen sind, können zusammenfassende Auswertungen nur für phasengleiche Abschnitte vorgenommen werden. Abb. 4 gibt eine tabellarische Zusammenfassung, wie lange die einzelnen Personen der Experimentalgruppe in den verschiedenen Stufen verweilt haben. Nach Auswahl von stufen-korrelierten Personen (siehe Abb.7, Experimentalgruppe 202 - 208, 210 - 212, 214, 217),  $n = 12$ ) zeigt diese "frisierter" Subgruppe aus der Stichprobe deutlich eine 4-Wochen-Rhythmizität entsprechend den von aussen vorgegebenen Bedingungen.

### 6.5.2. Resultate

Schon bei der grafischen Darstellung (Abb. 5, 6a, 6b, 6d, 6e) zeigt sich bei allen Kurven eine über die Zeit (Wochen) ausgesprochen grosse feste Periodizität von vier Wochen, die sich in entsprechenden Zeitreihenanalysen (Autokorrelation) durch hohe Werte für die Zeitabschnitte von Vier-Wochen-Intervallen äussern. Abb. 9b zeigt die Autokorrelation der wöchentlich erhobenen Token-Summenwerte über die Zeit. Der Korrelationskoeffizient zeigt ein Maximum bei 3-Wochen-Intervall, dann mit absteigenden Werten in etwa bei  $3 + 4$ ,  $3 + 7$  resp. grob gesagt  $3 + (n \times 4)$ , wobei  $n$  Integers (ganze Zahlen) zw. 1 - 6 sind. Eigentlich sollte der erste Peak bei  $t = 4$  Wochen zu erwarten sein, evtl. haben aber Streuungen zu einer solchen Verschiebung geführt (s. Abb. 9a, Autokorrelation über die Zeit eines einzelnen Probanden Nr. 203). Auch 10a zeigt den Punkteverlauf der einzelnen Probanden je für sich aufgezeichnet. Eine reine Schwarzweissdarstellung würde praktisch ein voll geschwärztes Band zeigen, deshalb sind hier die Mittelwerte von je 4 zufällig zusammengefassten Probanden dargestellt. Dabei ist die Streuung in der Zeitachse deutlich zu sehen. Auch die Punktezahlen (Maxima und Minima) schwanken bei den einzelnen Probanden sehr stark, was sich logischerweise auf die Korrelationskoeffizienten auswirkt. Die Beschränkung auf Auswertung von "frisierter" phasengleichen

Subgruppen (siehe weiter oben unter 6.5.1) ist nicht zulässig und entspräche einer Datenmanipulation im Nachhinein. Deswegen werden in der weiteren Folge auch noch sehr aufwendige Verfahren versucht oder mindestens theoretisch evaluiert.

Geprüft wurde ein Verfahren zur "Egalisierung" der Punktwerte der einzelnen Probanden. Statistisch ausgedrückt würde das bedeuten, die mittleren Amplitudenwerte der einzelnen Probanden zu egalisieren, das heisst, sie fallweise zu komprimieren oder expandieren. Zwar handelt es sich hier im Zeit/Amplitudenbereich um lineare Veränderungen, die theoretisch in der Zeitreihenanalyse (Frequenz/Intensität oder Spektraldarstellung) keine Aenderung der Grundfrequenz ergeben sollten (hier also des ersten Peaks), die Harmonischen werden aber verfälscht (siehe Abb. 9b, die "Peaks" resp. Wellenberge können auch als Harmonische eines 4 Wochenintervalles angesehen werden, deswegen weiter oben auch der Hinweis auf "n mal 4").

Ein Vergleich diverser Untergruppen der Tokens (Urinproben, Nemexineinnahme, Freizeitplan erstellen, Freizeitplan einhalten, Bett gemacht, Kleider/Toilettenartikel ordentlich aufbewahrt, in Haussitzung eigenes Problem einbringen, Urinprobe freiwillig, Urinprobe angeordnet) mit den Summenkurven zeigte bei keinem Probanden einen zeitlichen Zusammenhang.

Abb. 10b zeigt die gegenseitige Abhängigkeit von 6 Items in einer 6 x 6 Punkte - Matrix, berechnet stets in 1-Wochenintervallen (als "LAG " bezeichnet). Mathematisch wird hier eine Kreuzkorrelation zwischen allen 6 Items über die Zeitachse im Intervall von 1 Woche berechnet und dargestellt. Die graphische Darstellung - hier für Proband 203 - zeigt keine wesentlichen Abhängigkeiten, obwohl hier im Vergleich zu allen anderen durchgerechneten Versuchspersonen noch einige wenige "Balken" auftreten.

Diese Resultate zeigen auch, dass noch komplexere Analyseverfahren wie zB. multidimensionale Clusteranalyse keine schlüssigen Resultate liefern können, weil die möglichen "Punktewolken" in den einzelnen abgebildeten Dimensionen zu wenig dicht wären, um für Merkmalszusammenhänge herangezogen werden zu können. Als Beispiel mögen wir annehmen, 6 Items gegeneinander zu analysieren. Hierfür verwenden wir ein  $6 \times 5 \times 4 \times 3 \times 2 \times 1 = 720$ -dimensionales System. Bei vollständiger Korrelation von 2 items würden wir einen 18 Punkte Cluster in einer oder 2 "Ansichten" d.h. 2 dimensional Diagrammen, erhalten. Da aber schon in unserer einfacheren Analyse gezeigt werden konnte, dass kaum wesentliche Korrelationen vorhanden sind, sind wir weit weg von einer

solchen Konstellation. So erhielten wir vielleicht noch Ansammlungen mit 4 Punkten, die sicher nicht als Cluster bezeichnet werden können. Nähmen wir zusätzlich die Zeit als weitere Dimension hinzu, verlieren wir uns in einem ca. 5000 -dimensionalen System.

Ich hoffe, mit diesen zum Teil mässig vereinfachten Ausführungen gezeigt zu haben, dass es kaum möglich sein wird, hier weitere erfolgsversprechende Analysen vorzunehmen.

### **6.5.3 Diskussion der Token-Economy-Resultate:**

Bei der Token-Economy-Auswertung gibt es keine Vergleichsgruppe und die Resultate zeigen hier deutlich, dass die Verhaltensweisen der Dingi-Insassen von äusseren Faktoren, vor allem vom vorgegebenen Zeitplan, "gesteuert" werden konnten. Dies darf wohl für diese Personengruppe, bei der ein Sich-Einordnen allgemein schwerfällt, als Erfolg in Richtung "angepasstes, kontrolliertes Verhalten" angesehen werden. Mangels Vergleichsgruppe sind weitergehende Rückschlüsse nicht möglich. Wie weiter oben gezeigt wurde, konnten keine wesentlichen Zusammenhänge zwischen Token-Untergruppen festgestellt werden. Diese sind so heterogen (z.B. Urinproben auf Wunsch der Probanden möglich oder auf Anordnung, einige Probanden nahmen Nemexin) und zusätzlich durch Einzelausreisser in der Summenkurve (siehe weiter oben) schon "in der Zeitachse verschmiert", so dass wegen mangelnden Aussichten auf Resultate auf weitergehende Methoden als eine schon beschränkte multidimensionale Kreuzkorrelation, welche mit Einschränkungen als "Mini-Clusteranalyse" angesehen werden könnte, verzichtet werden kann.

### **6.6. Katamnestische Erhebungen**

Die Grundlagen hierzu sind vorbereitet. Wie die Resultate sein werden, ist noch offen, dies hängt z.T. auch davon ab, ob alle Probanden in der Schweiz bleiben werden und damit bezüglich weiterer Delinquenz überhaupt erfasst werden.

### **6.7. Zusammenfassung der Resultate, Diskussion und Ausblick**

In den beiden Fragebogen "FPI" (Freiburger Persönlichkeitsinventar) und "GT" (Giessen-Test) ergaben sich übereinstimmende Resultate, die alle für die Dingi-Gruppe (Experimentalgruppe) eine signifikante Veränderung in Richtung vermehrt sozialisierbar, allgemein bessere Stimmungslage und mehr Selbstbestimmtheit zeigen. Ebenfalls zutage getretene Hinweise für vermehrte

Aggressivität können als Begleiterscheinung des Dingi-Vollzuges betrachtet werden. Denn es handelt sich ja bei dieser Kohorte um schwer persönlichkeitsgestörte Individuen, bei welchen nun ein Anteil dieser Grundstörung ebenfalls möglicherweise in einem gewissen Sinn demaskiert worden ist. Wieweit allenfalls eine Interpretation in der Richtung möglich ist, dass diese Personen eventuell in Konfliktsituationen auch früher "Dampf ablassen" und damit zwar sozial unangenehm auffallen, aber dafür wenig delinquiren, wird sich zeigen müssen.

***Alle diese Befunde sprechen für eine bessere Ausgangssituation der Dingi-Gruppe für zukünftige Bewährung im Alltagsleben.***

Daraus kann abgeleitet werden, dass die Dingi-aner gegenüber den Kontrollpersonen aus dem Normalvollzug sicher eindeutig bessere Grundvoraussetzungen für die Sozialisation und Legalbewältigung haben. Daraus müsste theoretisch die zukünftige Delinquenz und auch Deliktschwere sowie Delikthäufigkeit zurückgehen.

In diesem Sinne kann festgestellt werden, dass die ***Anfangs-*** oder ***Startbedingungen für diese Hypothese*** mit dem Dingi-Vollzug ***erfüllt werden konnten. Eine weitergehende Aussage über die Haupthypothese gemäss 3.1*** (siehe auch unter 6.1) ***ist zur Zeit nicht möglich***. Sie kann erst aufgrund der Bewährung im dem Dingi-Vollzug nachfolgenden Lebensabschnitt anhand der katamnestic Erhebungen bestätigt oder verworfen werden.

Insgesamt zeigten sich auch Hinweise für eine nicht optimale Randomisierung der Subgruppen (Experimental- und Kontrollgruppe), dies wurde unter U-Fragebogen abgehandelt.

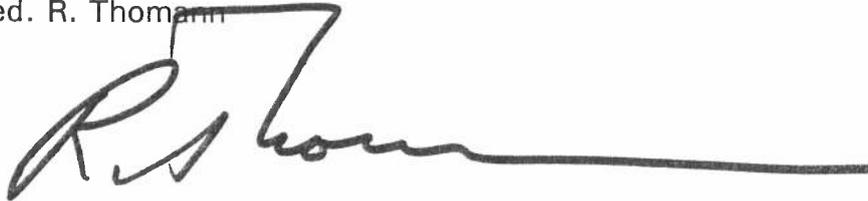
Bei der Ausarbeitung des Grundkonzeptes waren die genauen Auswertungsabläufe und auch die notwendigen Resultate, die für eine Annahme oder Verwerfung der Haupthypothese herangezogen werden, nicht formuliert worden. Entsprechend wird eine Antwort im Sinne "Hypothese erfüllt", d.h. Versuch gelungen respektive Experimental-Aussage positiv in diesem Sinne nicht möglich sein. Erschwerend kommt vielleicht noch hinzu, dass die genauen Auswertemechanismen erst spät anhand der schon festgelegten Abläufe erarbeitet werden mussten. Hierzu gehört auch die Token-Economy, welche zwar äusserst viele Rohdaten geliefert hat, jedoch lediglich eine einzige Aussage zugelassen hat.

**Ausblick:**

Die eigentliche Bewertung der Haupthypothese wird sich erst anhand der katamnesticen Nacherhebungen in einigen Jahren ergeben.

Königsfelden, 12. September 1998

Dr. med. R. Thomann

A handwritten signature in black ink, appearing to read 'R. Thomann', with a long horizontal stroke extending to the right.

**6.8. Dank:**

Den zwei Arztgehilfinnen Frau A.L. Notter und Frau J. Keusch möchte ich an dieser Stelle für die prompte Mithilfe bei einem Teil der Schreibearbeiten und Diagramm-Darstellungen danken, zumal die hierfür vorgesehene Hilfskraft zeitweilig unvorhergesehen nicht verfügbar war. Grosser Dank geht auch an den nicht genannt sein wollenden Leiter eines Rechenzentrums, ohne dessen Hilfe einige Auswertungen auf dem verfügbaren System nicht hätten durchgeführt werden können.

**6.9. Beilagen:**

Abbildungen 1a, 1b, 1c, 1d, 1e, 2a, 2b, 3, 4, 5, 6a, 6b, 6c, 6d, 6e, 7, 8a, 8b, 9a, 9b, 10a, 10b,

# DINGI FPI Test (Staninen)

## Kontrollgruppe N=14

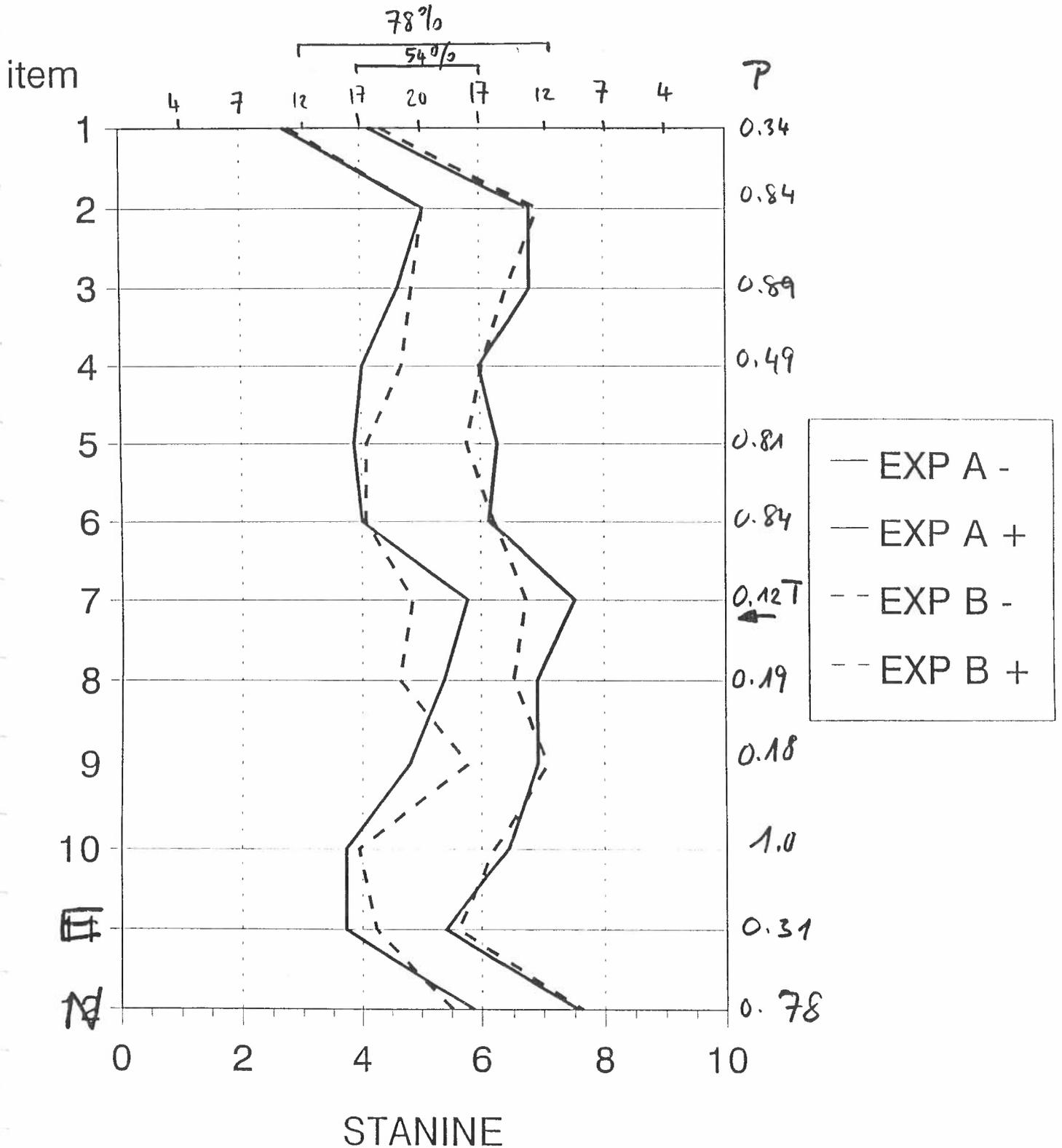
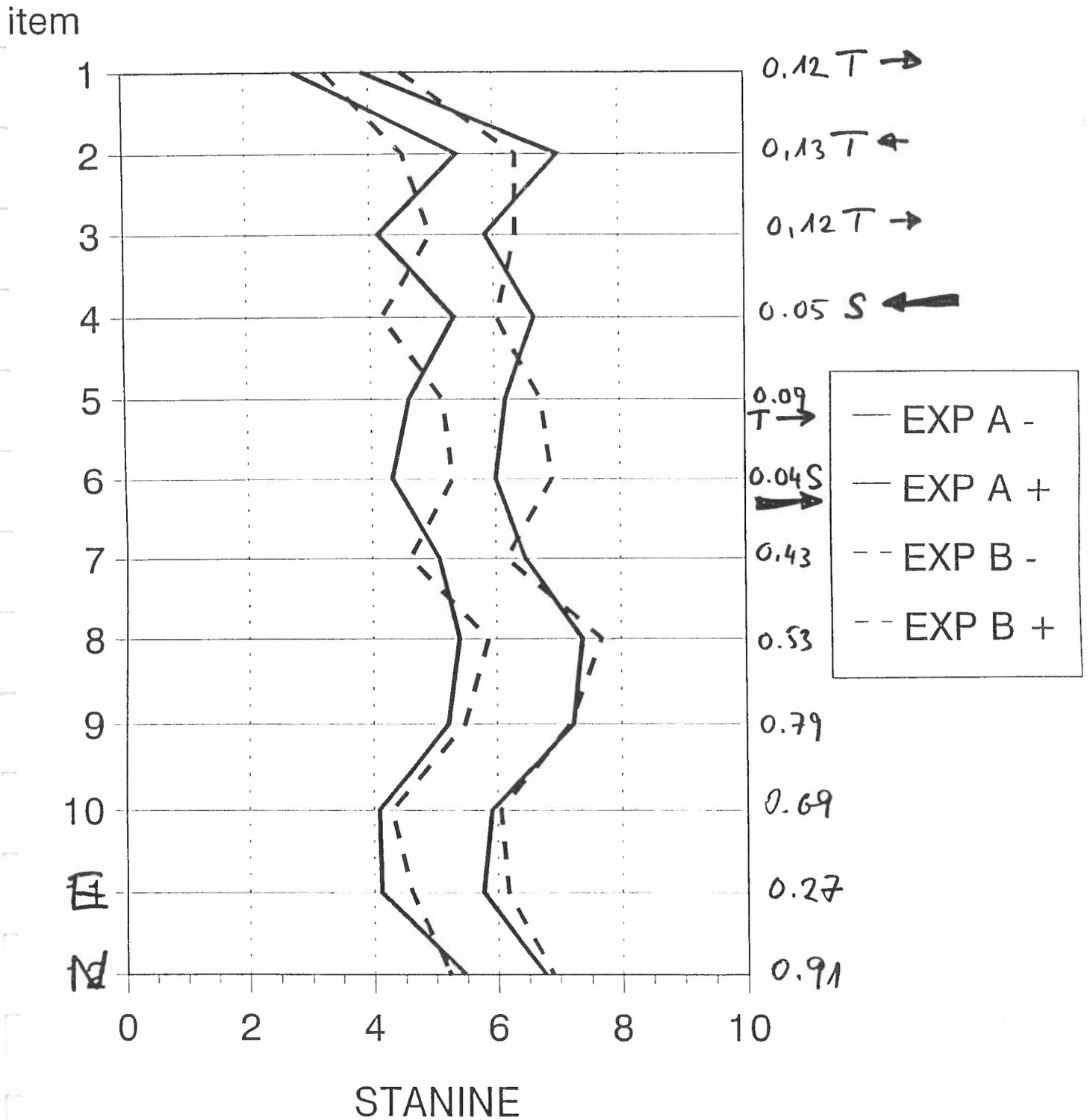


Abb. 1 a

# DINGI FPI Test (Staninen)

## Experimentalgruppe N=18



		FPI DINGI Experimentalgruppe 1.Messg.											
		Stani nen						altersgemacht					
		1	2	3	4	5	6	7	8	9	10	E	N
201	A	3	5	3	8	6	5	5	9	4	7	4	6
202	A	3	4	3	7	5	3	5	8	7	2	2	5
203	A	1	6	2	8	7	7	5	8	6	8	3	6
204	A	3	3	6	4	7	6	4	5	1	5	8	6
205	A	3	5	8	4	4	6	4	4	6	5	7	6
206	A	3	7	8	8	9	8	9	9	5	8	5	9
207	A	3	9	5	6	7	6	5	7	9	4	4	6
208	A	3	5	4	6	5	5	7	6	9	3	4	6
209	A	4	8	6	6	4	4	6	5	6	5	7	4
210	A	6	6	8	5	5	8	5	7	8	3	7	4
211	A	3	6	5	6	7	6	6	9	5	8	5	8
212	A	3	7	4	6	6	6	6	8	5	7	3	8
213	A	3	7	3	5	4	4	5	5	6	5	6	6
214	A	3	4	5	4	4	4	6	6	5	3	7	5
215	A	3	7	6	5	3	4	5	2	6	4	4	6
216	A	6	9	5	7	5	1	9	8	9	5	4	7
217	A	3	7	5	6	6	5	5	4	6	4	5	7
218	A	4	7	4	7	3	5	7	5	9	4	4	5
StdAbweichung		1.11	1.62	1.73	1.29	1.57	1.67	1.40	1.98	2.02	1.83	1.65	1.29
MW + 0.5 Stabw		3.89	7.03	5.87	6.65	6.17	6.00	6.48	7.38	7.23	5.91	5.77	6.75
Mittelwert		3.33	6.22	5.00	6.00	5.39	5.17	5.78	6.39	6.22	5.00	4.94	6.11
MW - 0.5 Stdabw		2.78	5.41	4.13	5.35	4.60	4.33	5.08	5.40	5.21	4.09	4.12	5.47

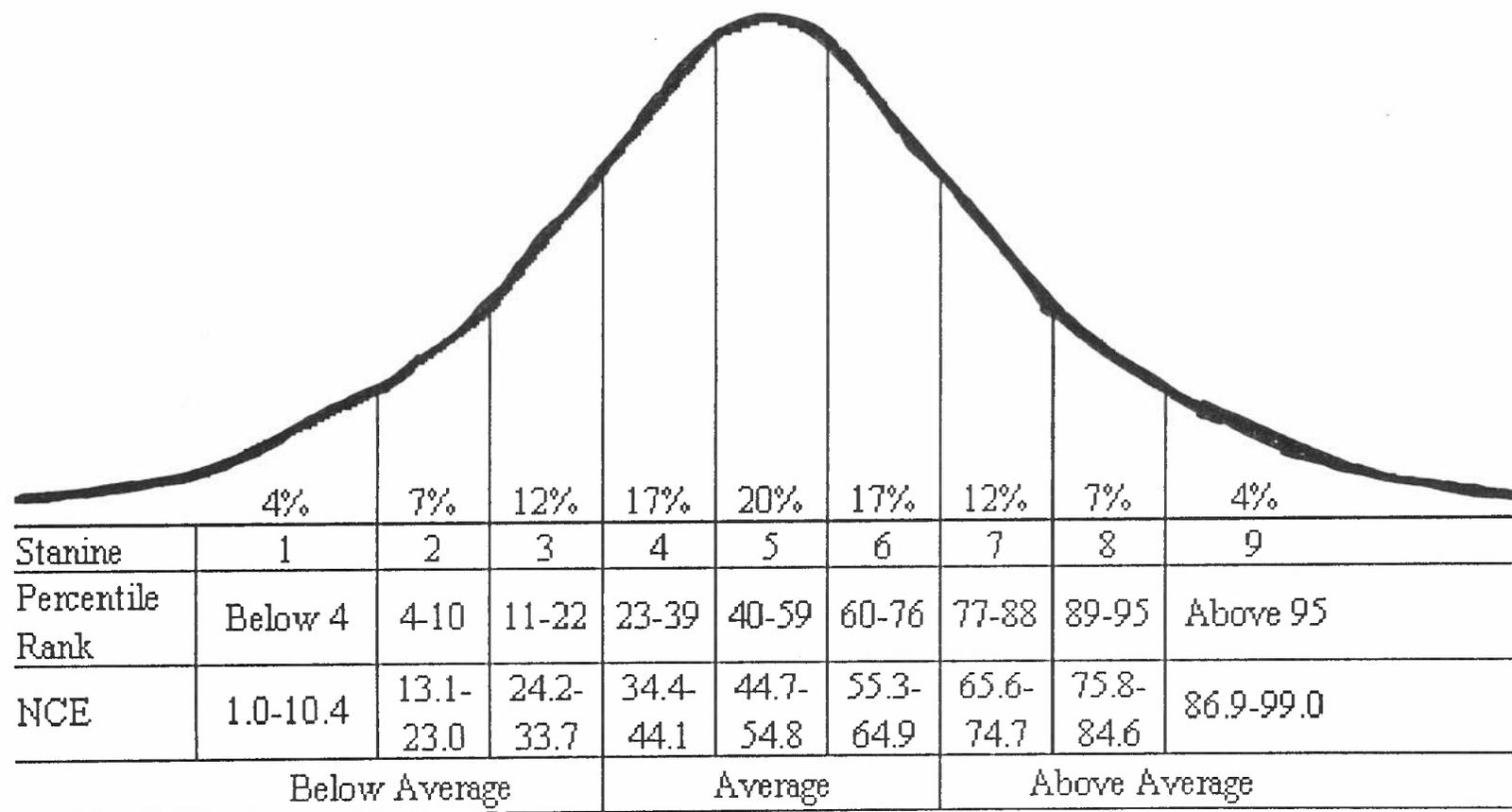
		FPI DINGI Experimentalgruppe 2.Messung											
201	B	4	5	5	5	4	6	4	7	6	7	5	5
202	B	6	5	7	4	5	6	4	5	6	4	4	5
203	B	2	6	4	8	6	9	6	7	8	7	4	7
204	B	3	4	8	4	9	7	4	6	4	4	7	6
205	B	5	6	6	5	7	6	6	9	8	4	5	8
206	B	3	4	6	7	9	8	7	8	5	9	5	9
207	B	4	7	7	7	7	7	7	8	8	6	5	7
208	B	4	5	5	7	6	4	2	9	9	4	3	8
209	B	4	9	5	8	5	4	6	8	4	7	8	6
210	B	7	4	8	4	4	9	5	8	6	7	7	6
211	B	3	5	7	2	8	6	6	8	5	4	6	5
212	B	4	7	7	5	6	6	6	5	6	7	5	7
213	B	3	4	4	3	4	5	5	6	6	5	8	5
214	B	3	3	5	4	4	4	9	6	3	2	4	6
215	B	5	6	5	3	5	5	4	2	8	3	5	2
216	B	2	2	5	4	7	8	4	4	6	5	8	4
217	B	4	7	5	4	6	6	5	8	7	4	5	5
218	B	4	9	3	8	5	4	7	8	9	4	3	8
StdAbweichg.		1.24	1.83	1.37	1.85	1.58	1.59	1.57	1.81	1.70	1.77	1.57	1.65
MW + 0.5 Stabw		4.51	6.36	6.35	6.04	6.73	6.91	6.17	7.68	7.18	6.05	6.17	6.88
Mittelwert		3.89	5.44	5.67	5.11	5.94	6.11	5.39	6.78	6.33	5.17	5.39	6.06
MW - 0.5 Stdabw		3.27	4.53	4.98	4.18	5.15	5.31	4.60	5.87	5.48	4.28	4.60	5.23
T Test symmetr. paired		0.12	0.13	0.12	0.05	0.09	0.04	0.43	0.53	0.79	0.69	0.27	0.91

Abb. 1 c (Staninenwerte)

Dingi FPI Kontrollgruppe 1. Messung												
Stanine Maenner Altersgemacht												
		1	2	3	4	5	6	7	8	9	10	E
	S	S	S	S	S	S	S	S	S	S	S	S
101	A	3	7	5	7	9	7	8	9	4	9	3
102	A	3	7	2	6	4	4	9	5	6	5	2
103	A	5	7	7	4	4	6	7	7	6	5	5
104	A	3	4	4	7	6	4	9	7	7	9	4
105	A	3	5	3	6	7	7	6	5	3	8	4
106	A	3	5	7	1	6	7	8	6	8	4	6
107	A	4	4	8	3	5	4	7	8	7	5	6
108	A	1	6	5	5	2	6	5	6	1	5	6
109	A	3	6	8	4	7	4	5	4	6	2	6
110	A	3	7	4	7	3	5	5	4	9	2	4
111	A	3	7	8	7	4	5	4	6	8	6	5
112	A	2	2	4	6	9	9	9	5	5	8	5
113	A	5	8	6	5	1	2	6	8	6	1	1
114	A	7	8	9	2	4	1	5	6	6	2	7
STABW		4.15	5.06	4.63	4.02	3.87	4.01	5.77	5.39	4.80	3.72	3.72
rt		3.43	5.93	5.71	5.00	5.07	5.07	6.64	6.14	5.86	5.07	4.57
STABW		2.70	6.79	6.80	5.98	6.27	6.14	7.51	6.90	6.91	6.42	5.42
		1.45	1.73	2.16	1.96	2.40	2.13	1.74	1.51	2.11	2.70	1.70

FPI KO												
						Dingi	Dingi	K				
101	B	3	6	7	6	4	6	7	8	7	4	4
102	B	3	8	5	4	3	2	5	4	6	6	4
103	B	5	4	5	6	3	6	2	5	7	4	5
104	B	3	6	2	7	4	5	7	6	5	7	3
105	B	3	5	5	6	6	7	7	8	5	9	5
106	B	3	7	6	5	7	7	7	6	6	4	5
107	B	3	4	6	7	4	4	7	9	7	6	6
108	B	2	7	5	6	5	6	6	5	4	8	5
109	B	3	5	9	4	6	4	5	2	7	3	5
110	B	3	6	6	6	6	7	7	6	9	5	4
111	B	4	7	6	5	4	4	4	5	8	5	8
112	B	2	2	5	6	9	9	9	4	6	7	5
113	B	5	9	5	5	4	1	4	6	7	1	3
114	B	8	8	7	2	4	4	4	4	6	2	7
STABW		4.35	6.94	6.42	6.03	5.77	6.21	6.71	6.51	7.07	6.21	5.62
rt		3.57	6.00	5.64	5.36	4.93	5.14	5.79	5.57	6.43	5.07	4.93
STABW		2.79	5.06	4.87	4.69	4.09	4.07	4.86	4.64	5.79	3.94	4.24
		1.55	1.88	1.55	1.34	1.69	2.14	1.85	1.87	1.28	2.27	1.38
ymm.		0.34	0.84	0.89	0.49	0.81	0.84	0.12	0.19	0.18	1.00	0.31

Abb. 1 d (Staninen -Werte)



*A Normal Distribution of Stanines, Percentile Ranks,*

Abb. 1 e

# GT Test Experimentalgruppe N = 18

*/* = Erstmessung      */* = Zweitmessung

Prozentrang      0,1   0,6   2,3   6,7   15,9   30,9   50   30,9   15,9   6,7   2,3   0,6   0,1

T      4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 42 44 46 48 50 52 54 56 58 60 62 64 66 68 70 72 74 76 78 80 82 84 86 88 90 92 94

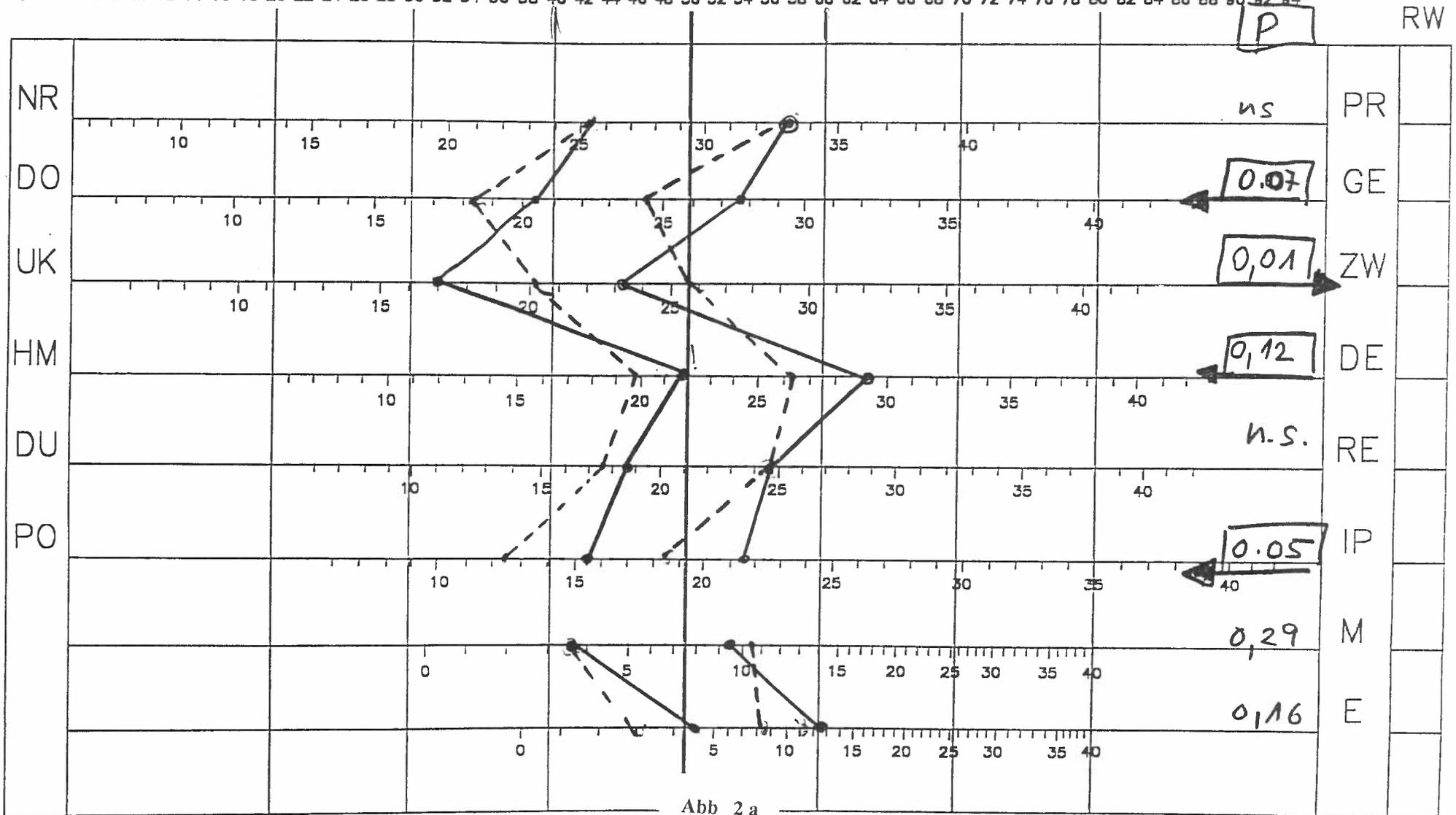


Abb 2 a

# GT Test Kontrollgruppe N = 14

 = Erstmessung     
  = Zweitmessung

Prozentrang     0,1   0,6   2,3   6,7   15,9   30,9   50   30,9   15,9   6,7   2,3   0,6   0,1  
 T     4   6   8   10   12   14   16   18   20   22   24   26   28   30   32   34   36   38   40   42   44   46   48   50   52   54   56   58   60   62   64   66   68   70   72   74   76   78   80   82   84   86   88   90   92   94

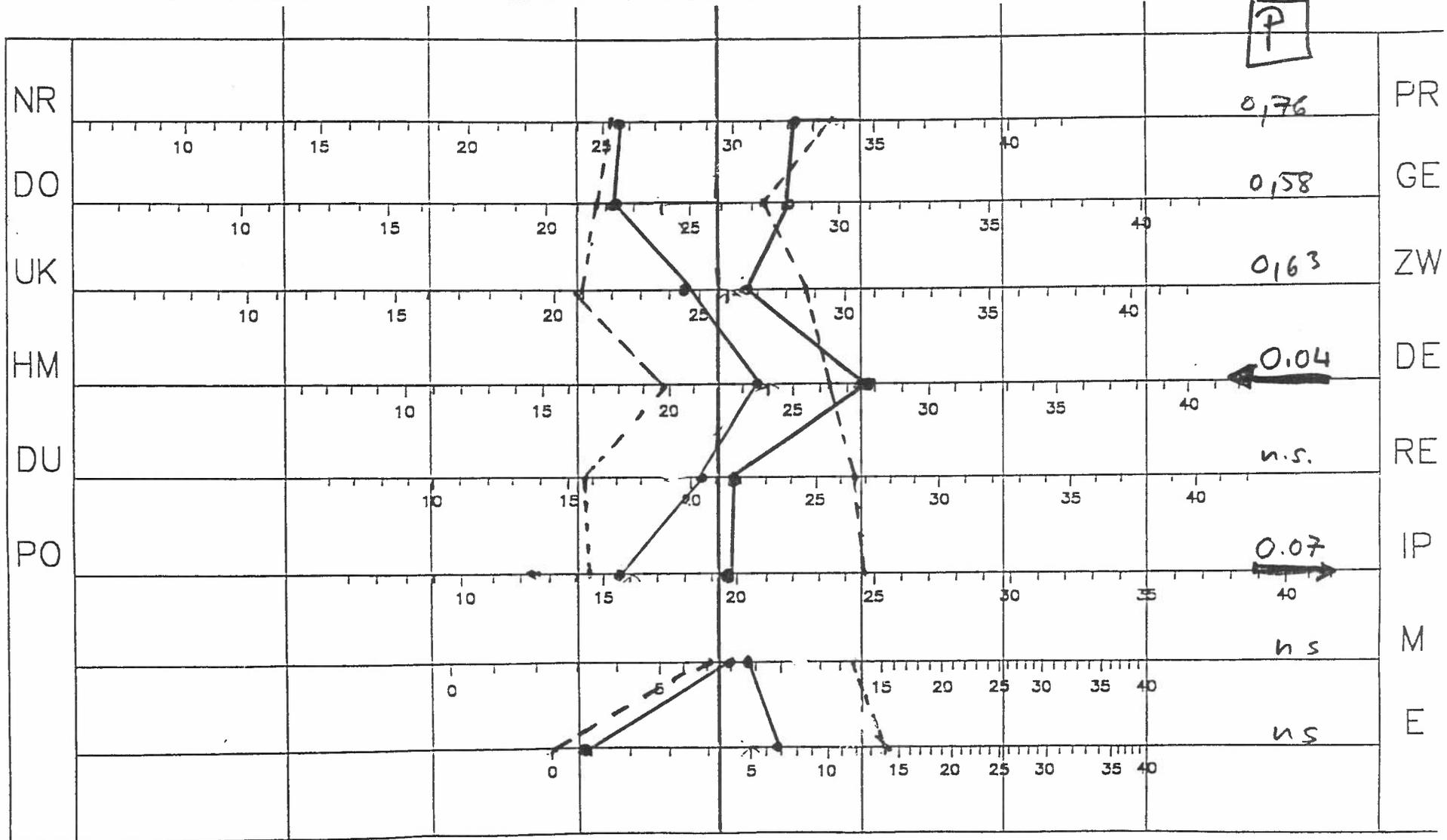
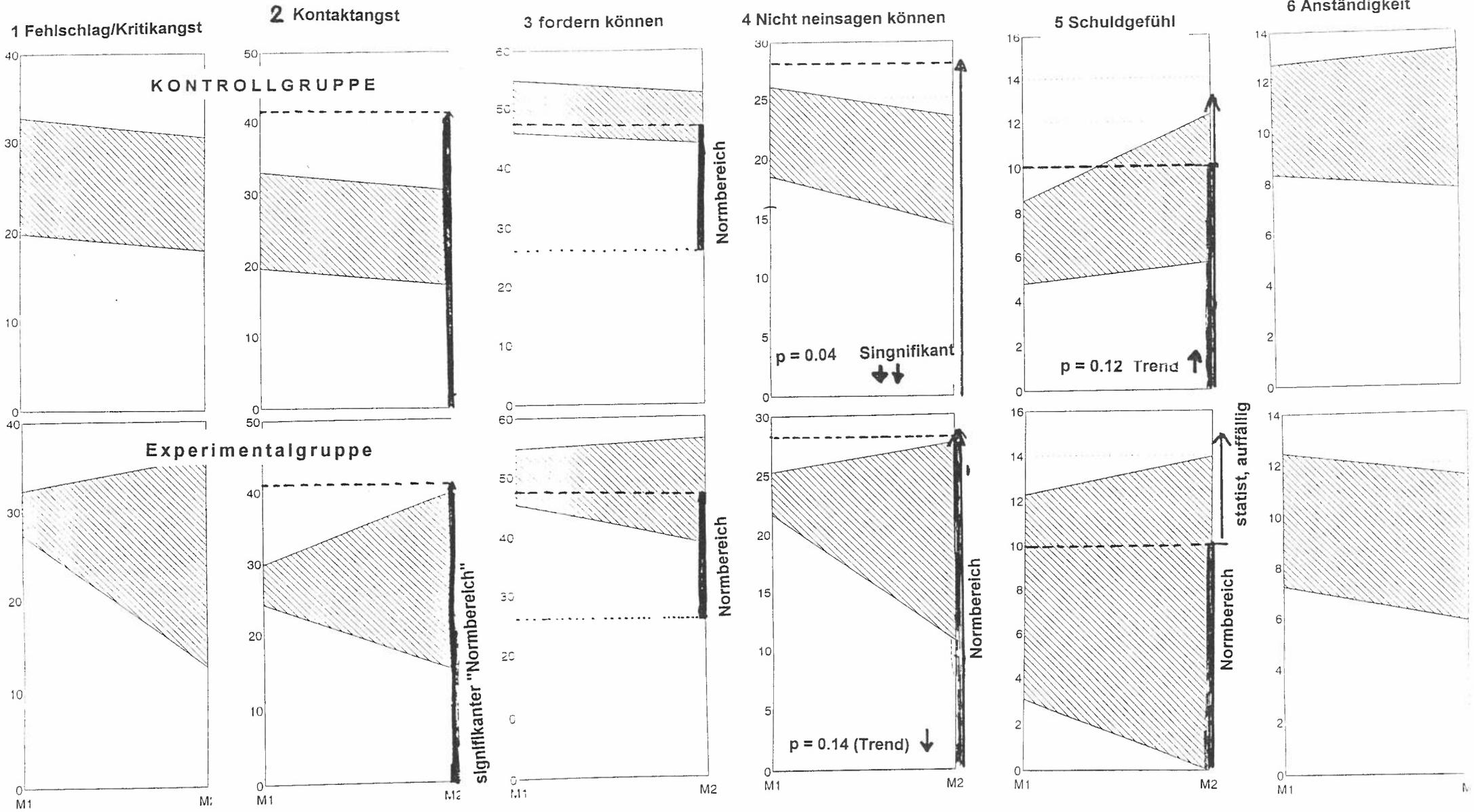


Abb 2 b

# Dingi U Fragebogen



M1 = Erstmessung M2 Zweitmessung

Abb. 3

# Dauer pro Stufe und Exp.Person

## DINGI Tokens Experimentalgruppe

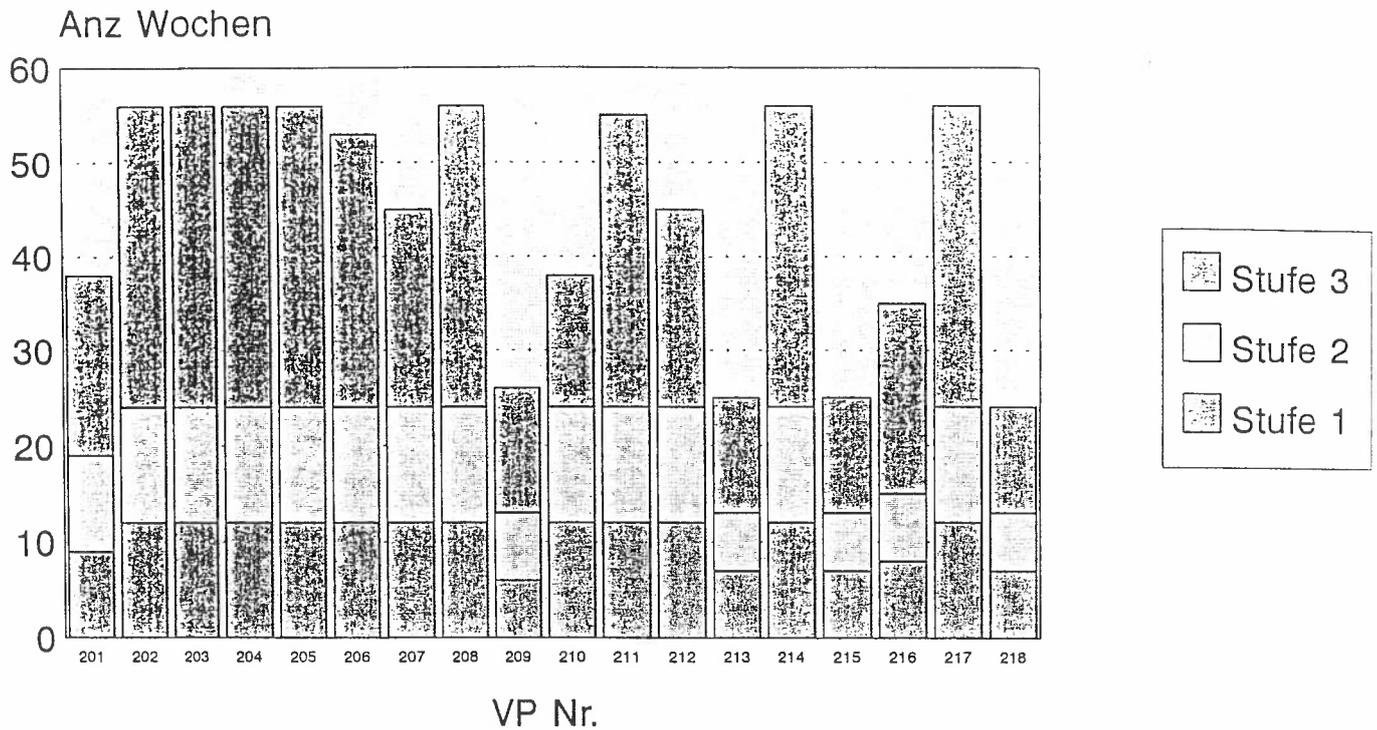


Abbildung 4

7.8.97 9.8.97 korr

# Token Summe vs Zeit

## Dingi Experimentalgruppe

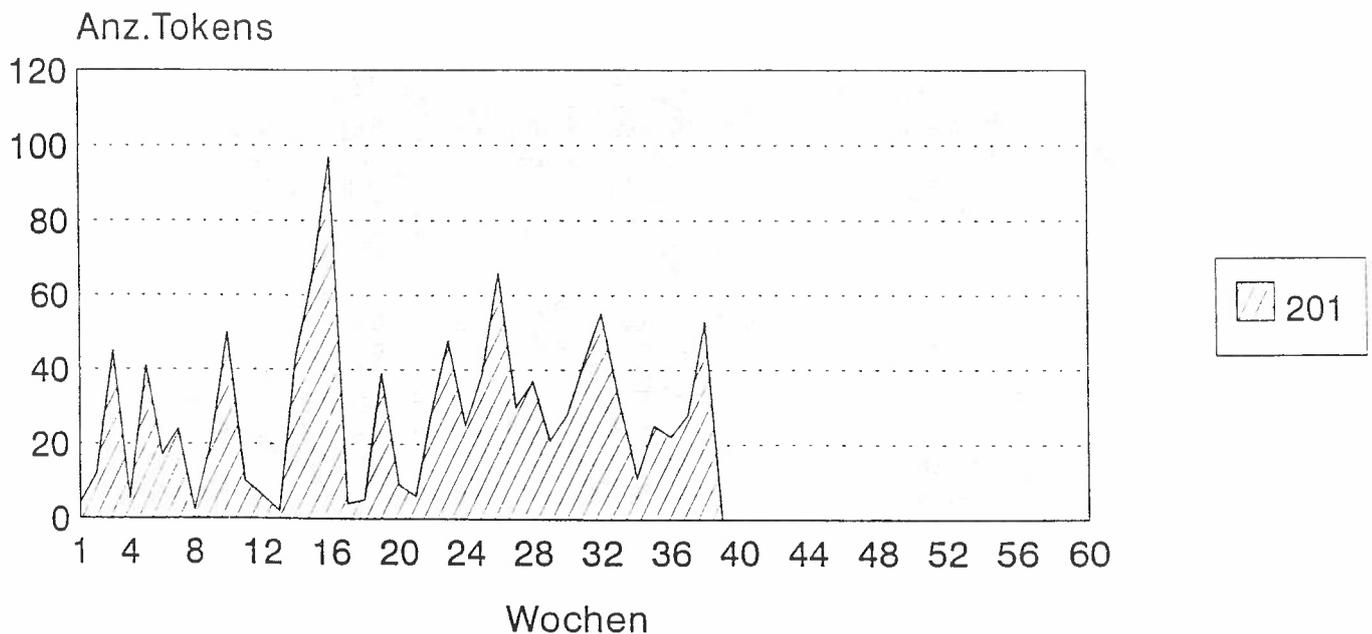
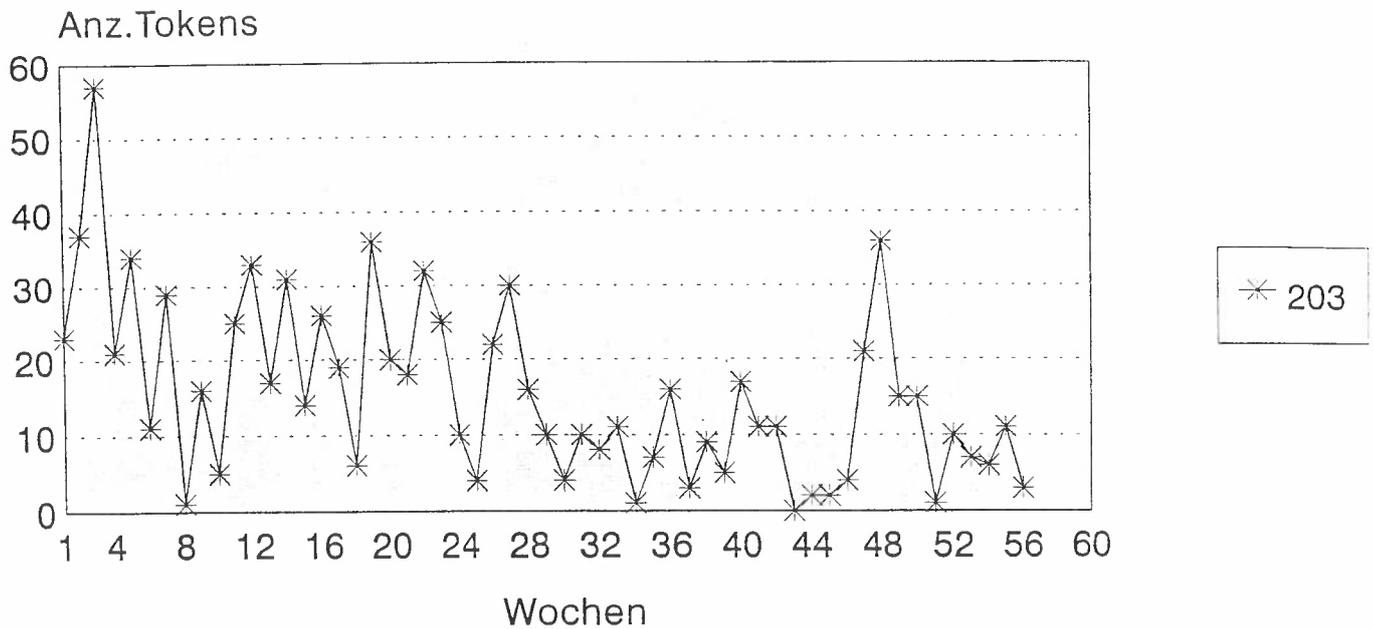


Abbildung 5

# Token Summe vs Zeit

Dingi Experimentalgruppe

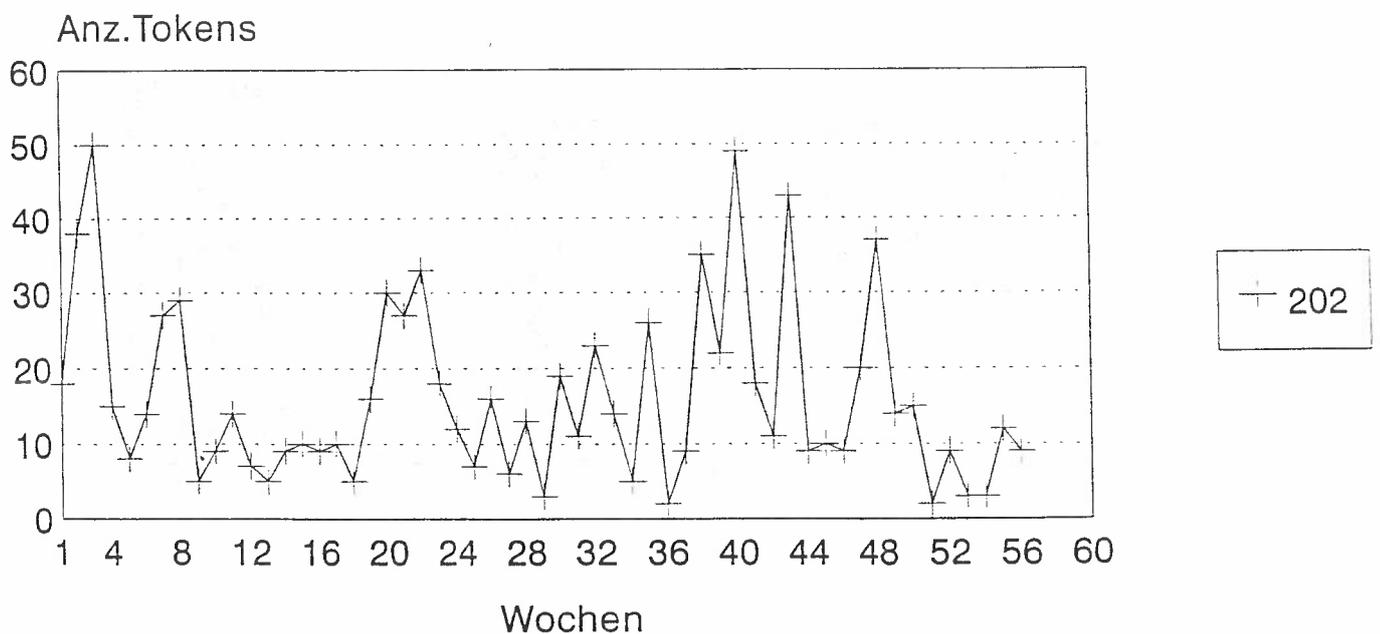


/TH/Dez 97

Abbildung 6A

# Token Summe vs Zeit

Dingi Experimentalgruppe



/TH/Dez 97

Abbildung 6B

# Token Summe vs Zeit

Dingi Experimentalgruppe

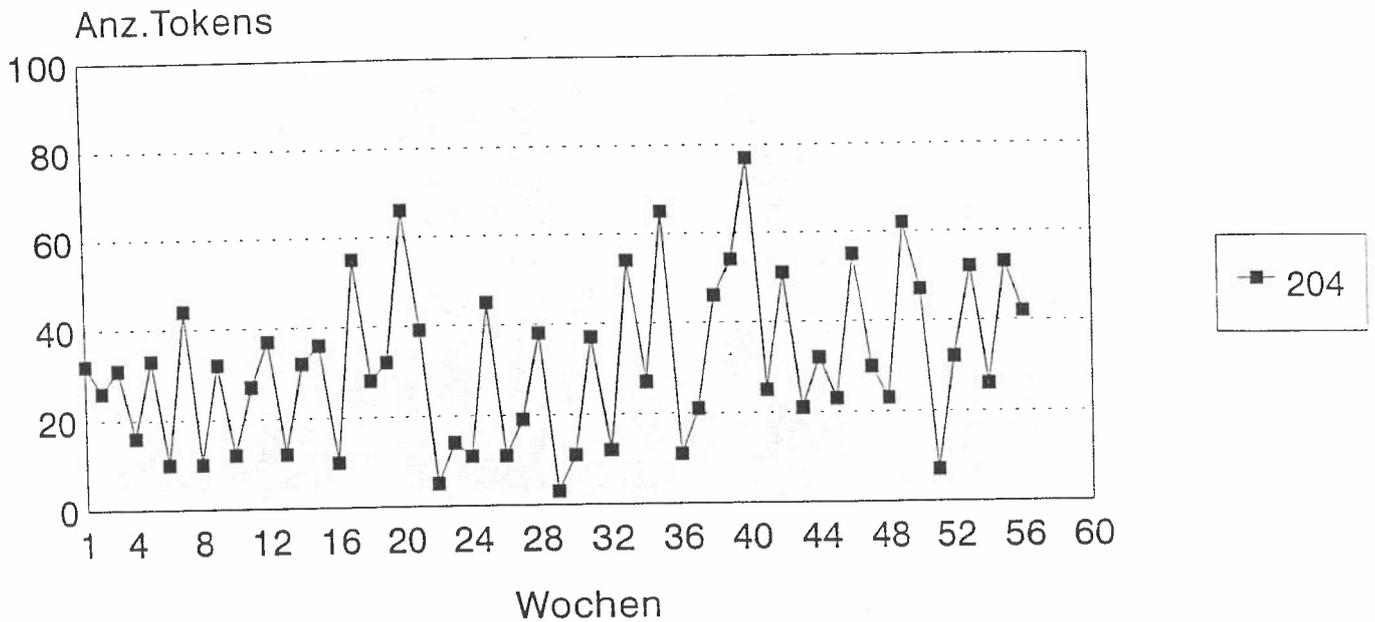


Abbildung 6C

/TH/Dez 97

# Token Summe vs Zeit

Dingi Experimentalgruppe

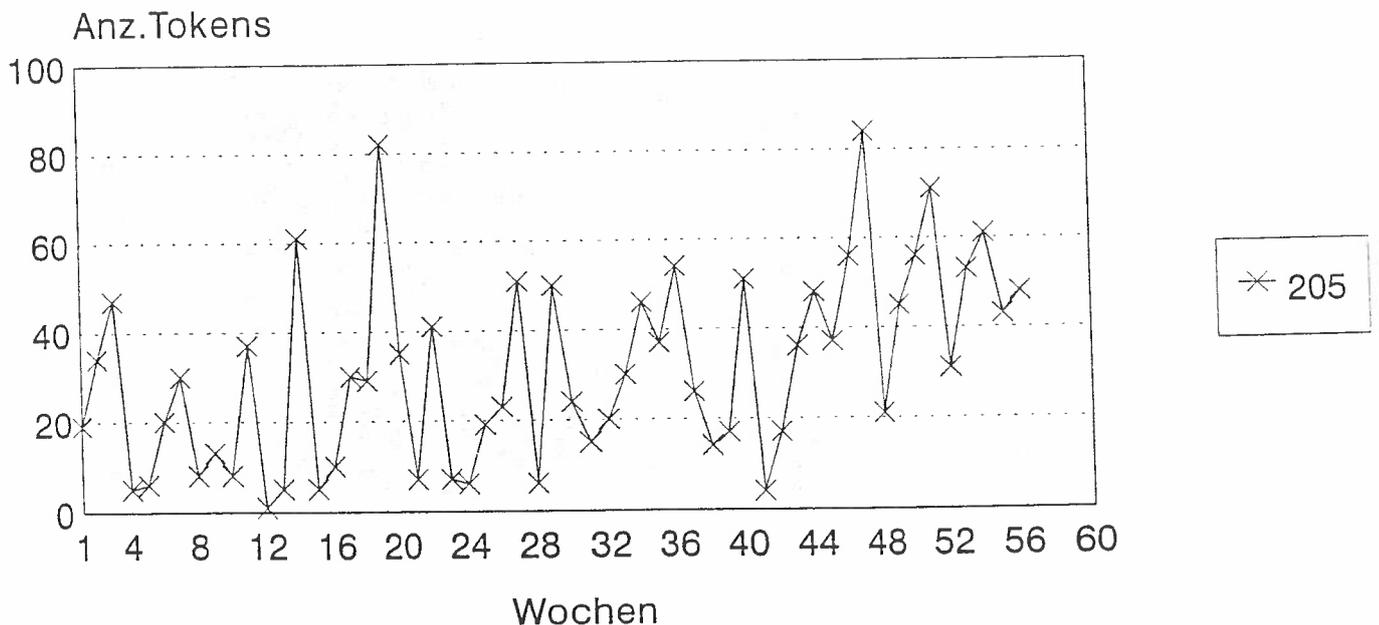
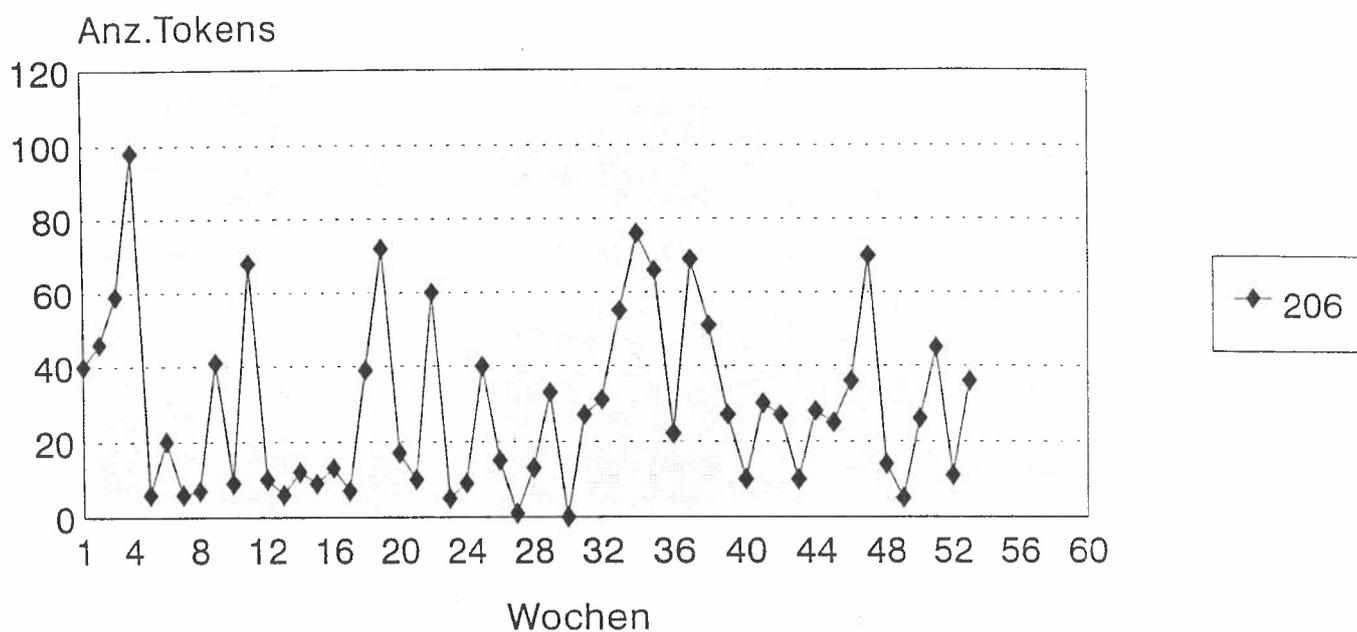


Abbildung 6D

/TH/Dez 97

# Token Summe vs Zeit

Dingi Experimentalgruppe



/TH/Dez 97

Abbildung 6E

## Tokens MW/StdAbw etc. Koordiniert nach Stufe

DINGI Exp (202-208,210-212,214,217)

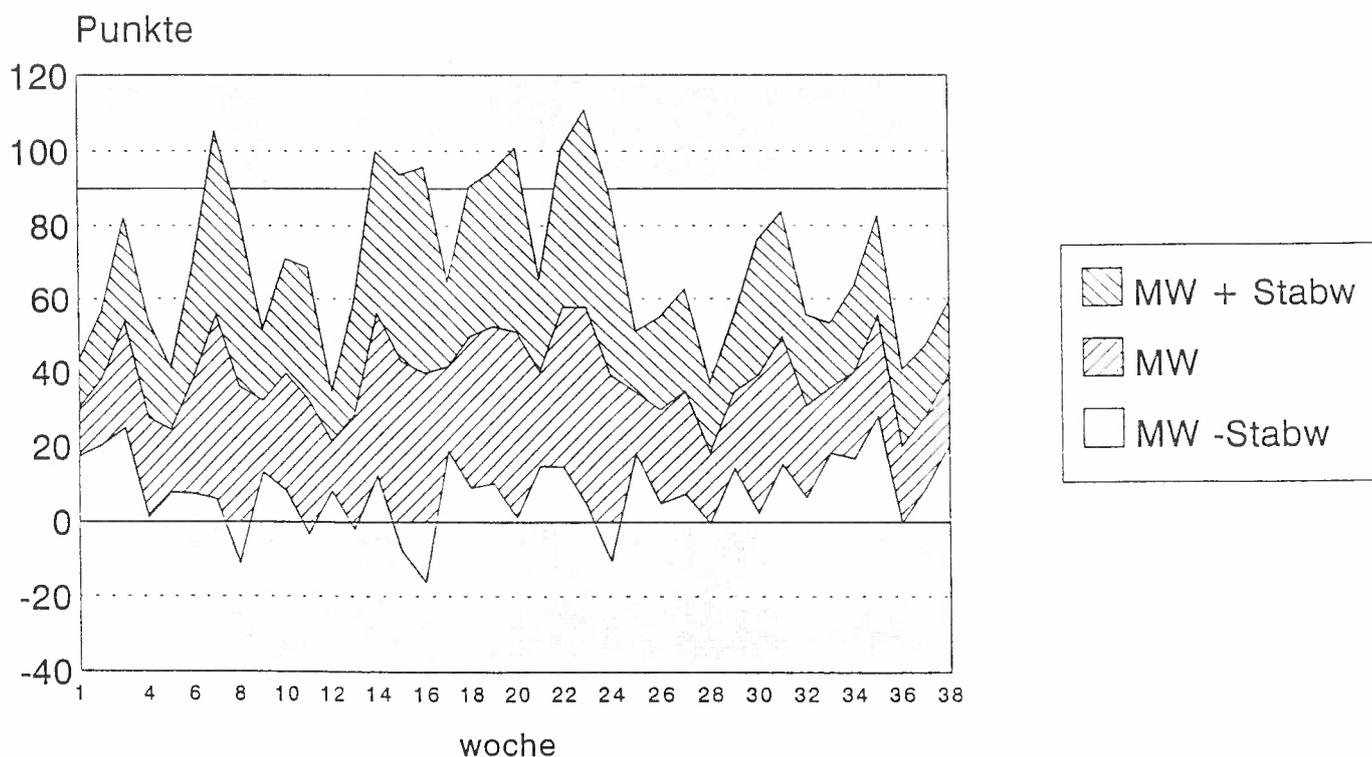


Abbildung 7

# Uebergang Woche 15,16,17

## DINGI Exp. Tokens

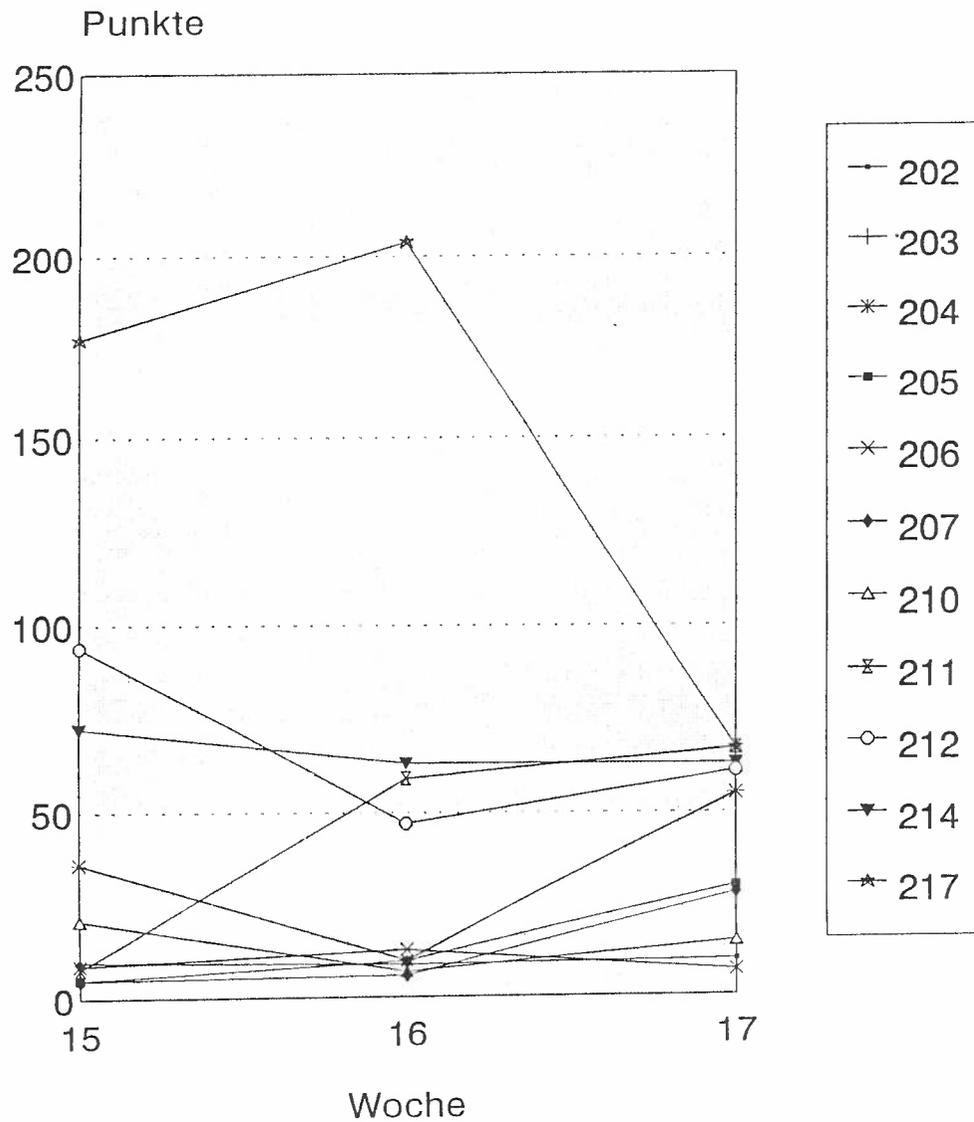


Abbildung 8A

Vers. 19.8.97

# Uebergang Woche 11,12,13

## DINGI Experimental

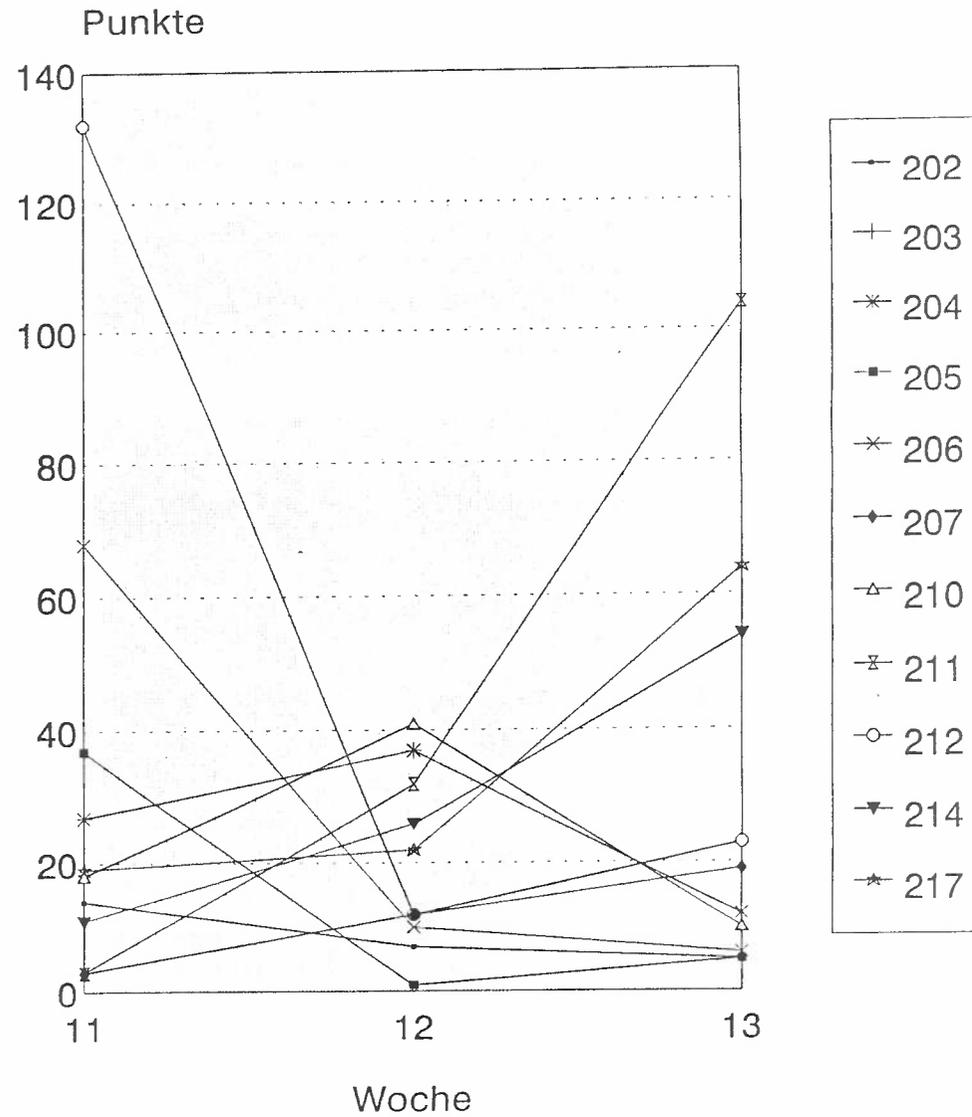
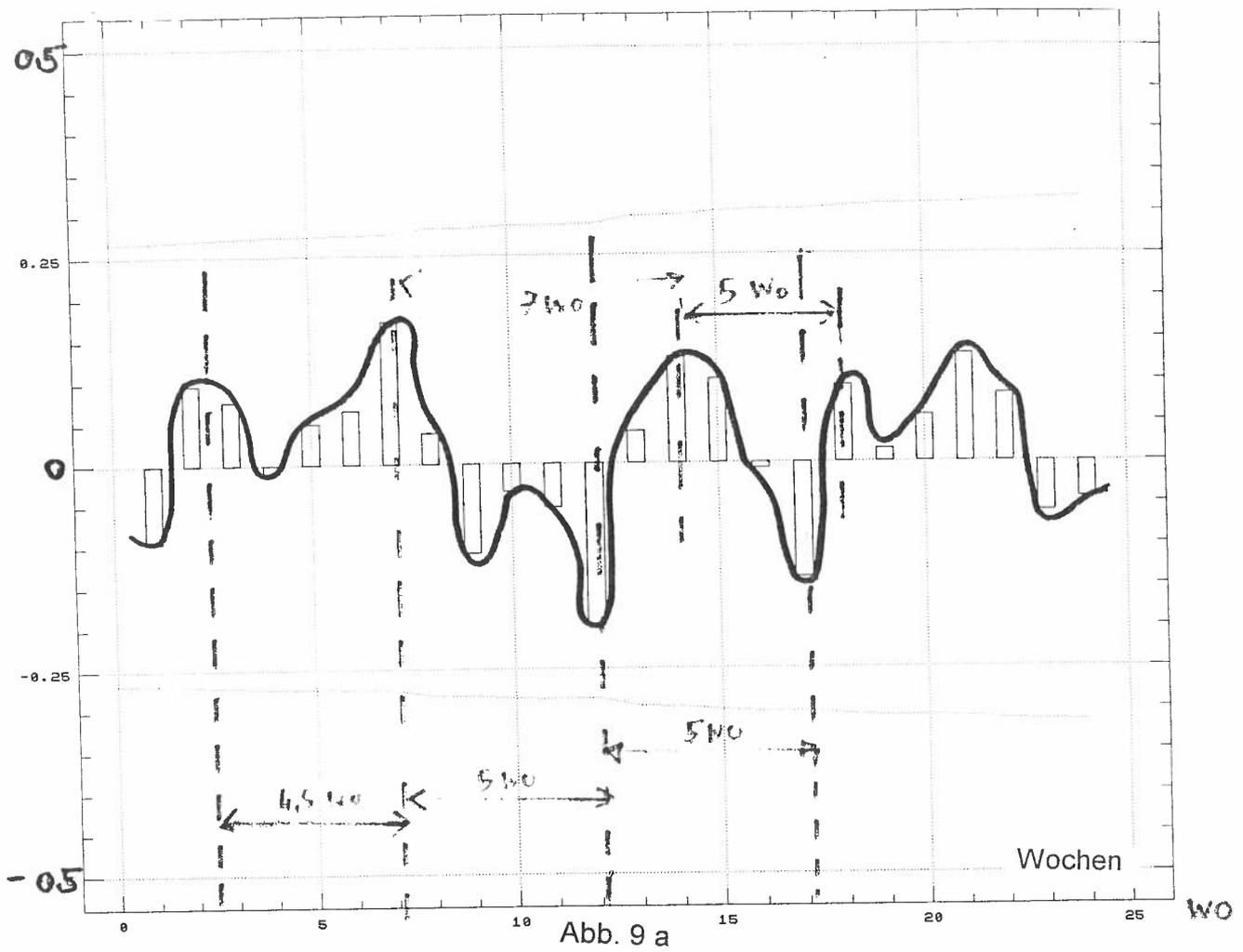
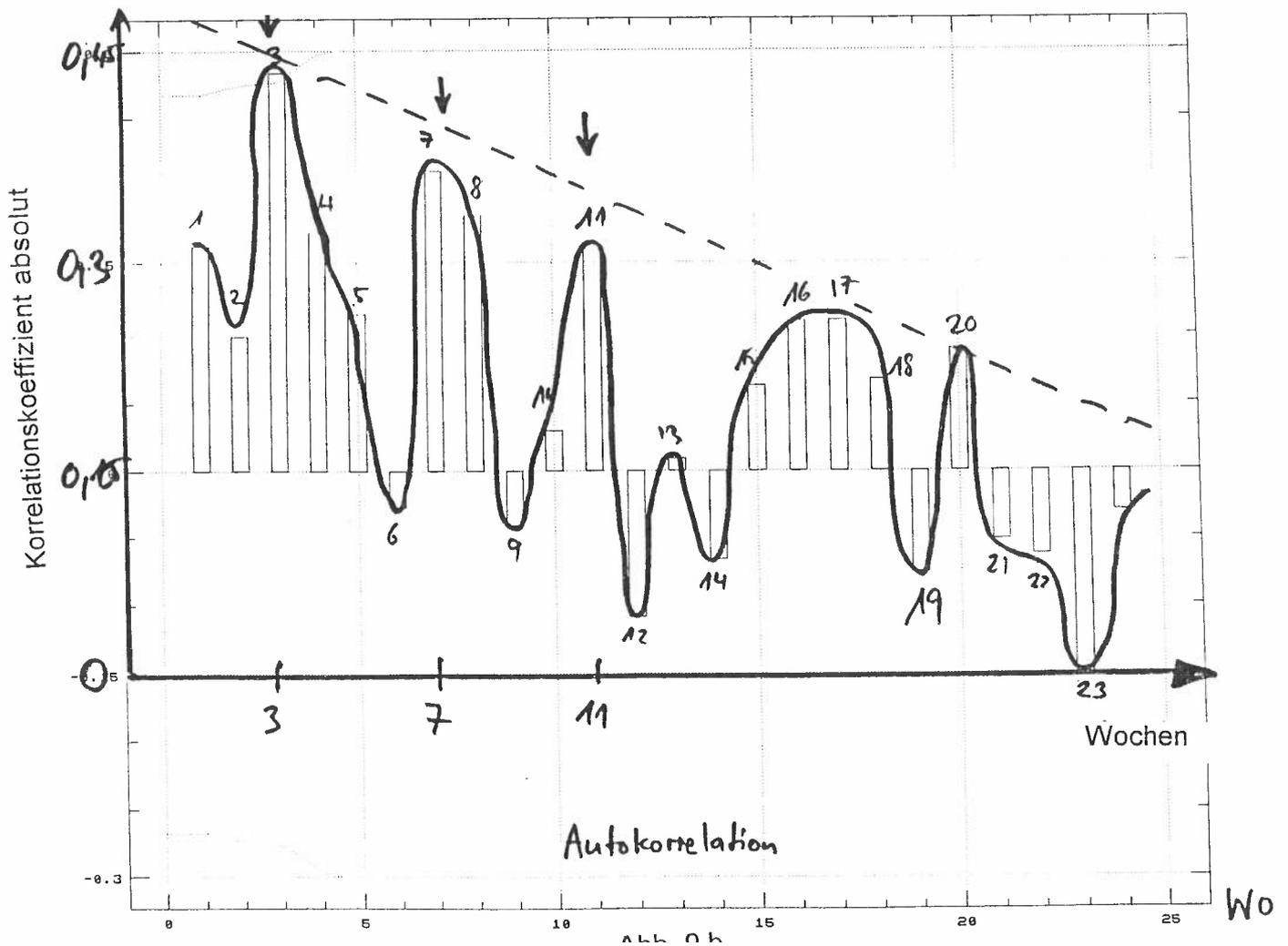


Abbildung 8B

/TH Vers. 8/97 rev 11/97



Tokeneconomy: Autokorrelation der Summen über Zeit



# Tokeneconomy Verteilung der Untergruppen

TOK.U202

--- TOK.U203

TOK.U205

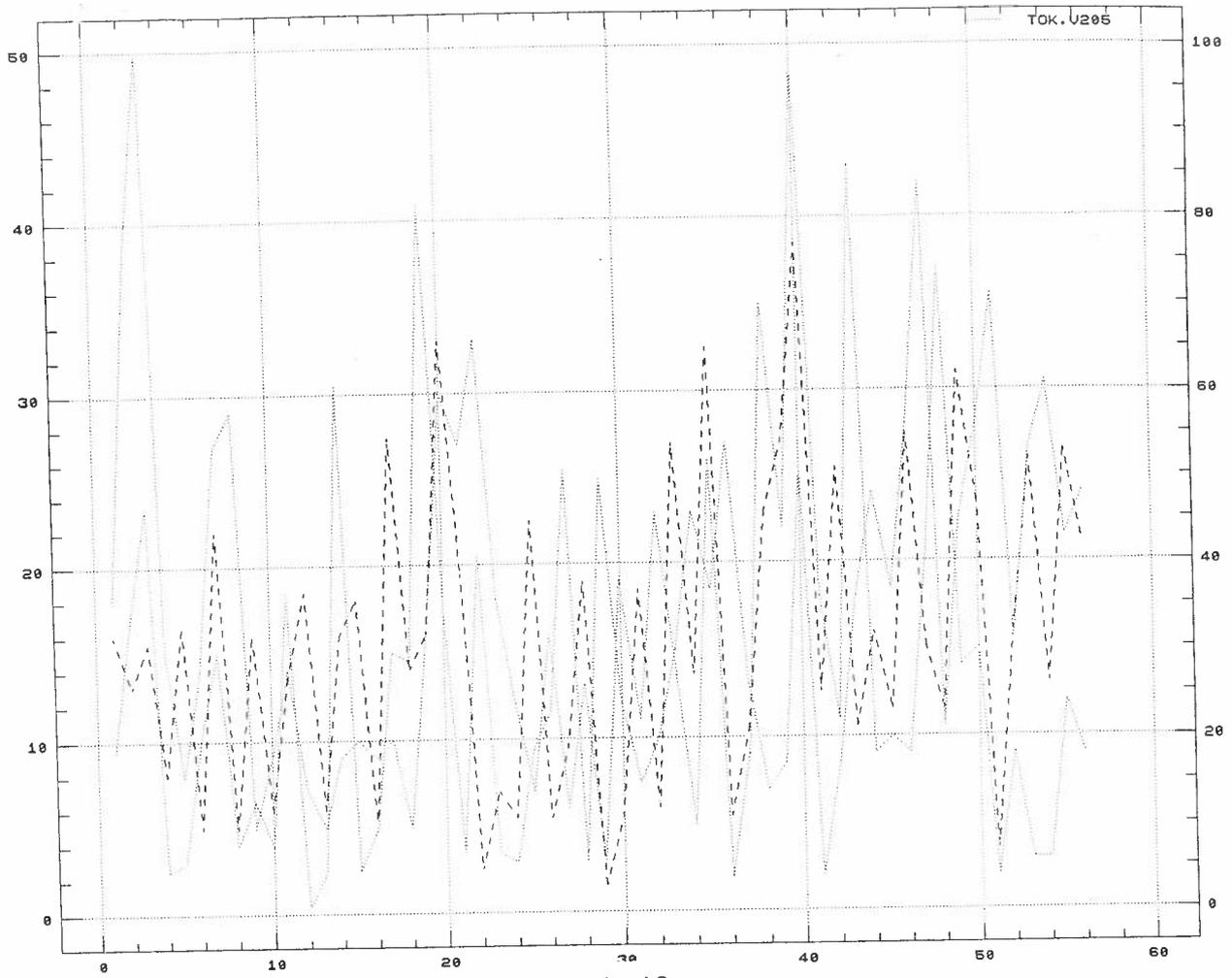


Abb. 10 a

## Crosskorrelation der Tokeneconomy Untergruppen vs. Summenkurve über Zeit

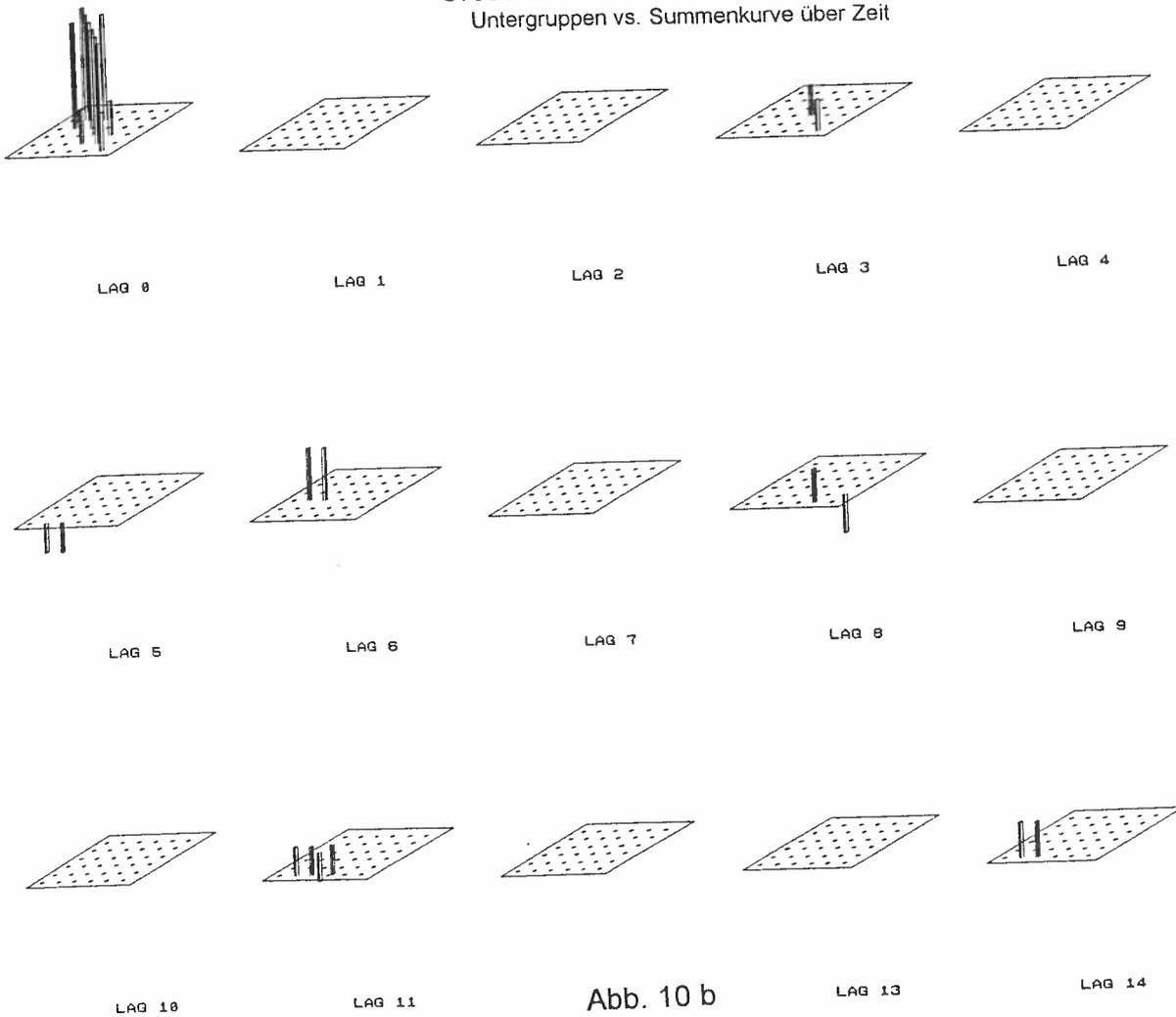


Abb. 10 b