Thought and action

Mobach, M.P.

Publication date:
1999

Chapter 4 Validating our Methods[11]

4.1. Introduction

In the first part of *phase one*, this pilot study, a test of several methods was made in community pharmacy practice. Since there were no validated methods for collecting data on the issue of this thesis, we had to formulate and test some methods ourselves. Previously, in the modelled world of **chapter 3**, three pharmacy mixes were described analytically. It was also argued that *thought* related to the perceived importance of activities and *action* related to the actual use of activities. Since we will now start to step in the real world, we have decided to use the term 'action' rather than 'activity'. Although actions relate to observations (for example, describing the importance and performance of actions of a pharmacy manager), we decided to use this terminology also in the design of questionnaires. Both *thought* and *action* are described in relation to the three pharmacy mixes mentioned before. A total of seven methods were tested. The main purpose of the pilot study is selection of methods. The pilot study seeks to produce two methods which can analyze both features best for a large sample. The large sample will be used in the second and next part of *phase one* of the study, which is described in **chapter 5**. In addition, in this pilot it was studied whether there was a correspondence between *thought* and *action* of pharmacy managers. This additional material is used at most to sketch the pharmacy mixes in practice. This pilot study seeks to find out which of the methods used can best be applied to identify correspondence of *thought* and *action* of pharmacy managers for a large sample. The results of the pilot study cannot be used to make generalizations valid for the population of Dutch community pharmacies; only 24 Dutch community pharmacies were studied. The next part of *phase one* will be made in exactly the same way. This next part will be studied with the methods which are selected here.

selected mix *thought*	**product mix** e.g. *ranking*: 1. product 2. process 3. customer		**process mix** e.g. *ranking*: 1. process 2. customer 3. product		**customer mix** e.g. *ranking*: 1. customer 2. product 3. process	
selected mix *action* (max)	corr	*non*-corr	corr	*non*-corr	corr	*non*-corr
product-related (92)	**90**	20	20	50	50	20
process-related (111)	50	**90**	**90**	**90**	20	40
customer-related (100)	20	50	50	20	**90**	50
remaining (69)	20	20	20	20	20	**60**
missing	192	192	192	192	192	202
total possible score (372)	**372**	**372**	**372**	**372**	**372**	**372**

Table 4.1. An example of a correspondence and *non*-correspondence of *thought* and *action*.

[11] Parts of this chapter were published in: Mobach MP, Werf JJ van der, Tromp TFJ. APOM-project: a first study of pharmacy organization and management. *Pharm World Sci* 1998; 20(5): 219-224.

Within this current pilot study, correspondence and absence of correspondence of *thought* and *action* were studied. **Table 4.1.** shows some examples of an ideal correspondence (corr) and *non*-correspondence (*non*-corr) of *thought* and *action* of pharmacy managers. The columns show correspondence and *non*-correspondence of *thought* and *action*. Per pharmacy mix, an example of a ranking for *thought* is presented. The examples are: product mix (product-process-customer) in the second and third columns, process mix (process-customer-product) in the fourth and fifth columns, and customer mix (customer-product-process) in the sixth and seventh columns. In general, the rows show the maximum scores of pharmacy-mix-related actions. The categories are: product (92) in the second row, process (111) in the third row, customer (100) in the fourth row, remaining (69) in the fifth row, and, missing are presented in the sixth row. There is a total maximum of 372 possible actions in the sixth row. Ideally, the ranking of the selected pharmacy mix (*thought*) should correspond with the main actions of pharmacy organization (*action*). How was this analyzed? An example.

In a test interview the pharmacy manager was asked to prioritize issues relating to the product mix, the process mix, and the customer mix. The pharmacy manager ranked the product mix as the most important issue, the process mix as of second importance, and the customer mix as of final importance (compare second upper left cell in **table 4.1.**; *ranking* product-process-customer). This outcome relates to *thought*. In another test interview the pharmacy manager was asked for the actions of pharmacy organization relating to the product mix, the process mix, the customer mix and remaining actions. The main actions observed were product-related actions. Second best were process-related actions, and third best were customer-related actions (compare second column **table 4.1.**; actions 90-50-20-20-192). This outcome relates to *action*.

The second column in **table 4.1.** shows that the ranking for *thought* (product-process-customer) corresponds with the ranking for *action* (product[90]-process[50]-customer[20]). The ranking for *thought* product-process-customer is the same as the ranking for *action*. The pharmacy manager mainly concentrates on actions like the specific standards the medicine should meet, the minimization of risks, and the minimization of error occurrence etc. In this example *thought* and *action* of the pharmacy manager corresponded (corr). The third column of **table 4.1.** shows an example of *non*-correspondence (*non*-corr). The ranking product-process-customer in *thought* is not at all related to the ranking for *action* (product[20]-process[90]-customer[50]), being process-customer-product.

Another example. A pharmacy manager selects the ranking process-customer-product for *thought*. In the fourth column of **table 4.1.** the same ranking is observed for *action* (product[20]-process[90]-customer[50]). In this example, *thought* and *action* of the pharmacy manager corresponded. The pharmacy manager mainly concentrates on actions such as the way in which the dispensing of the prescription is organized and to what extent it is organized profitably and efficiently etc. However, the fifth column of **table 4.1.** shows another example of *non*-correspondence. The ranking process-customer-product in *thought* is not completely related with the ranking for *action* (product[50]-process[90]-customer[20]), being process-product-customer. Although the first ranking for *thought* and *action* corresponds, namely product, neither other rankings corresponded. This situation is defined as *non*-correspondence. This study is in general looking for complete correspondence (in all three positions of the rankings) between *thought* and *action*.

4.2. Design of the pilot study

The first questionnaire was formulated in December 1994. The final design of the questionnaire was completed in May 1995. During this period, comments were received from the supervisors, stichting VNA, SAL Apotheken, and field experts. Two categories of field experts were interviewed: pharmacy managers and experts working in the community pharmacists' sector. It was assumed that the feedback of the field experts would contribute substantially to the content of the questionnaire. Seven pharmacy managers and four field experts gave comments on the questionnaire. The pharmacy managers were selected randomly. The four field experts were active in the field of regulation (Inspectie Volksgezondheid), health-insurance (Zorgverzekeraars Nederland), patient interests (NP/CF) and software systems (PharmaPartners); all with respect to community pharmacy. The questionnaire was adapted in accordance with the comments. The adaption mainly related to the content of the questions. Per question the number of issues was completed, mostly added. In addition, the form of the questions was adapted. Within the pre-pilot phase of the study a five-point scale was tested (not important-very important) with seven pharmacy managers. The pharmacy managers regarded most of the issues as being very important. As a result hardly any difference between the scores of various subjects was observed in the outcome of the five-point scale. To avoid this problem a three-point ordinal scale was introduced; in order to force the pharmacy manager to make a priority per question.

After completion of the final design in May 1995, the questionnaire was applied to 24 community pharmacies from May through to August 1995. Twelve pharmacies were selected by stichting VNA and SAL Apotheken, and were connected with stichting VNA and SAL Apotheken (VNA/SAL pharmacies). Twelve additional pharmacies were randomly selected among pharmacies in the twelve Dutch provinces; one pharmacy out of each province (additional pharmacies). The relatively large number of 12 VNA/SAL pharmacies (case numbers 13-24) were selected because of the interest of stichting VNA and SAL Apotheken in this study. The VNA/SAL pharmacies were encouraged by the top of the organization; nonresponse would hardly be expected. The additional pharmacies (case numbers 1-12) were selected to receive a more general picture of community pharmacies in the Netherlands.

At each pharmacy a questionnaire was used to get a *profile* of the pharmacy organization. In addition, at each pharmacy a combination of seven questionnaires was used with respect to *thought* and *action* of pharmacy managers (**table 4.2.**). The methods for *thought* consisted of the 123-method A, the 123-method B, the card method and the drawing method. The methods for *action* consisted of the method for the pharmacy manager, the pharmacy personnel (second pharmacist and assistant pharmacists) and the method for the researcher.The combination of the methods and the pharmacies was made randomly, with the precondition that 50% of the selected pharmacies were VNA/SAL pharmacies, and 50% of the selected pharmacies were additional pharmacies, for all individual methods. As a result, all four individual methods related to *thought* consisted of 6 VNA/SAL pharmacies and 6 additional pharmacies. For example, **table 4.2.** shows that 123-method A consisted of 6 VNA/SAL pharmacies (case 18, 23, 14, 16, 17 and 19) in the first column and 6 additional pharmacies (case 6, 9, 3, 11, 1 and 4) in the second column. In addition, all three methods relating to action consisted of 4 VNA/SAL pharmacies and 4 additional pharmacies. At twelve pharmacies the questionnaires with respect to *action* were sent after the questionnaire with respect to *action* were received by the researcher, vice versa for the remaining twelve pharmacies.

thought	**123-method-A**		**123-method-B**	
action	**cards**	**drawing**	**cards**	**drawing**
manager	VNA/SAL (case 18, 23)	additional (case 6, 9)	additional (case 7, 10)	VNA/SAL (case 13, 22)
manager & staff	VNA/SAL (case 14, 16)	additional (case 3, 11)	additional (case 2, 5)	VNA/SAL (case 15, 21)
manager & researcher	VNA/SAL (case 17, 19)	additional (case 1, 4)	additional (case 8, 12)	VNA/SAL (case 20, 24)

Table 4.2. The combination of methods and pharmacies studied.

4.3. Methods

All statistical procedures were made in SPSS 6.01 for Windows from September through to December 1995. Roughly three statistical methods were used: frequency distribution, Friedman test and cluster analysis for *profile*, *thought* and *action* respectively. A total number of 21.464 items were used for the pilot study. The statistical methods used for *profile*, *thought* and *action* will be discussed below.

One method was used to study the ***profile*** of pharmacy organization at all 24 pharmacies. Several subjects were analyzed with respect to general features of the pharmacy manager and of the pharmacy organization. Features of the pharmacy manager were Seniority, Sex, (In)dependence and Division of Time. Features of the pharmacy organization were Organizational Form, Cooperation, Location, Part-time and Full-time Personnel, Full-Time Equivalence (FTE) Pharmacists, FTE Other Personnel, Flow of Prescriptions, Flow of Patients, Turnover and Net Profit. Within this method a nominal scale was used (**table 4.3.**).

method	***profile***	***thought***				***action***		
quality		123-method A	123-method B	card method	drawing method	pharmacy manager	pharmacy personnel	researcher
form	survey	survey	survey	survey	survey	survey	survey	interview
question	1 out of k[12]	rank 3 out of 3	rank 3 out of 3	rank 4 out of 10	draw 3 positions	p[13] out of k	p out of k	p out of k
scale	nominal	ordinal	ordinal	ordinal		nominal→ ratio	nominal→ ratio	nominal→ ratio

Table 4.3. Qualities of the applied methods in the pilot study.

Four methods were tested to study the ***thought*** of the pharmacy manager. All methods with respect to *thought* were made by the pharmacy manager. At each pharmacy a questionnaire was sent in which pharmacy managers prioritized actions (123-methods). Additionally, two short methods, in which pharmacy managers prioritized their actions (card method and drawing method) were used, twelve pharmacies per short method.

All questions of the 123-methods consisted of three sub-questions with respect to the three pharmacy mixes. The sub-questions were ranked on a scale from 1 to 3, representing *important -less* important *-even less* important issues. The 123-methods consisted of the subjects Information, Administration, Automation, External Contacts, Facilities, Analysis, Organization of Labour, Personnel, Competence, Organization

[12] k relates to the number of possible options.

[13] p relates to the number of selected options.

Standards and Productivity. Firstly, for twelve pharmacies the method consisted of 28 questions in which the relation of the question to the pharmacy mix was mentioned (123-method-A). Secondly, for twelve pharmacies the method consisted of 45 questions in which the relation of the question to the pharmacy mix was *not* mentioned (123-method-B). Within these methods, an ordinal scale was used (**table 4.3.**).

In addition, two short methods were used. The first short method consisted of two sets of 10 cards, each card describing a subject related to a specific pharmacy mix. Each set of cards was divided into three groups: *important*, *not important* and *rest*. Per set 3 or 4 cards were to be selected in the category *important*. Within this category the cards were ranked from 1 to 4, representing most to least important, respectively. Within this method an ordinal scale was used (**table 4.3.**). The second short method consisted of a drawing in the form of a triangle of the three pharmacy mixes. Three positions were ranked by marks in the drawing, representing an angular point, moving along the axe and a diagonal-line endpoint (**figure 4.1.**). The pharmacy manager located the pharmacy organization in the drawing.

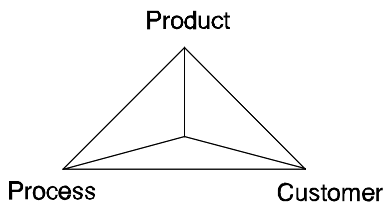

Figure 4.1. The triangle representing three pharmacy mixes.

Three methods were tested to study the ***action*** of the pharmacy manager. All three methods consisted of the subjects Information, Administration, Automation, External Contacts, Facilities, Analysis, Organization of Labour, Personnel, Competence, Organization Standards and Productivity. The pharmacy manager of each pharmacy was sent a questionnaire. The method for the pharmacy manager was the main instrument to measure *action*. It was expected, however, that the problem of social desirability could be at stake. Social desirability means that the answers of respondents are also determined by what they think the researcher will value as a good answer (Swanborn 1987), or the other way around, that there are certain facts or events that respondents rather would not report accurately (Fowler 1984). In the context of this study this would mean that the pharmacy manager could have been able to select questions in accordance with correspondence. In order to avoid this problem, other methods for the pharmacy personnel and for the researcher were used as well. In this situation, if correspondence in *thought* and *action* was observed, it would still be possible to test the realism of this outcome with the outcome of two additional methods, namely, for the pharmacy personnel and for the researcher. But if correspondence and *non*-correspondence were measured with the instrument for the pharmacy manager, it was assumed that no further comparisons were required for the purpose of the next part of *phase one*. At eight pharmacies only the questionnaires for the pharmacy manager were used. At eight pharmacies the personnel was also sent a questionnaire, and at eight pharmacies the researcher also interviewed the pharmacy manager (**table 4.2.**). All the questionnaires consisted of binary questions (true/false) related to the actions at the pharmacy organization. Within this method a nominal scale was used, and, in addition, the questions were rescaled via a count variable to ratio scale (**table 4.3.**).

4.3.1. Profile method

With the questionnaire for ***profile*** a total number of 1.393 items was used in the statistical processing. The data were used for a frequency distribution. Issues related to the pharmacy manager (Seniority, Sex, (In)dependence and Division of Time) and issues related to the pharmacy organization (Organizational Form, Cooperation, Location, Part-time and Full-time Personnel, FTE Pharmacists, FTE Other Personnel, Flow of Prescriptions, Flow of Patients, Turnover and Net Profit) were in general used for a general picture of the participating pharmacies. In addition, the data were used to produce information per pharmacy.

4.3.2. Thought methods

With the questionnaire for ***thought*** a total number of 2.904 items was used in the statistical processing. The data were mainly processed with a Friedman test. The Friedman test processing consisted of 123-method A, 123-method B and card method. The test was used to produce a mean ranking per pharmacy. The drawing method was not processed statistically, but will be presented as a pictogram.

The data were measured with the use of an ordinal scale. We asked ourselves how the results of this ordinal scale could be calculated? Could a mean ranking per case be calculated? Within 123-method A and 123-method B all observations were ranked, 1 being important, 2 less important, and 3 even less important. Within the ordinal scale of the 123-methods the distance between 1, 2 and 3 has no meaning. However, amongst statistical experts a controversy exists with respect to treating ordinal scales as interval scales and, for example, calculating a mean. The field is divided into liberals and conservatives (Knapp 1990). The liberals argue that although they do not have a true interval scale, they regard the differences between categories A and B and C as equal. In the methods used in this current study the distance between priority 1 and priority 2 would then be 1. They argue that it makes little difference whether an ordinal scale is treated as an interval scale. For example, liberals could propose just a calculation of the mean. Conservatives would counter that researchers have demonstrated very strange results when using means, standard deviations, and Pearson's r's with ordinal scales. Within this current study the conservative view was applied, although we also calculated some ordinary means. A distribution-free test procedure was used; the Friedman test. If the observations can be ranked, the Friedman test is a very powerful analytic tool. It calculates a mean ranking and, additionally, calculates a p-value. The Friedman test can be applied to a comparison between more than two groups and related samples (Slotboom 1987). Sometimes, when sample sizes are very small, there is no alternative to some form of distribution-free test of significance (Rosenthal and Rosnow

1991). According to Kolstoe (1973) the Friedman test was developed for use with correlated groups. The Friedman test requires that different observations of one individual (or matched group) must be capable of being ordered (ranked).

The ranking applied within the 123-methods per three sub-questions was dependent. For example, the number 1 could only be used once per three sub-questions. The outcome of the other two sub-questions depended on the outcome of the initial ranking (for example, value 1), then only the values 2 and 3 were still available. The Friedman test takes dependence between 1, 2 and 3 into account with the calculation of the mean ranking. However, then a problem arises. Within the Friedman test, independence between cases is required. For example, independency is assumed if different pharmacy managers were asked to rank three subjects. As a result, a mean ranking of one question for all different cases could be calculated within the Friedman test. However within this study, a mean ranking was required per case for every question. Within one case each question was answered by one pharmacy manager. Within the 123-method A and B only one pharmacy manager ranked several subjects. Could the Friedman test still be applied? Dependence between the various questions was present, since one person made the test. However, within this study independence within one case was assumed if the pattern in answers was the same for different sets of questions in one case. Independence of the various sets of questions of one pharmacy manager had to be proved. It was assumed that the questions presented were a sample out of a universe of N possible questions. Primarily, each single question was independent of all other questions. This means that any set of questions should produce merely the same pattern in answers and no learning effect in the test should be present. However within the 123-methods, many questions were answered by the same pharmacy manager. It was tested in a binomial-test procedure whether the pattern in the answers to the first 33% of the questions was merely the same as the pattern of the last 33% of questions. If the same pattern of answers was produced, absence of learning effect and independence of the questions was assumed.

The test of independence was made by computing the z-value for the normal approximation of the difference between two proportions.[14] *The computed z_c was compared with the table of cumulative standardized normal distribution F(z) (Harnett & Murphy 1986) with a level of significance of α=0.05. It was tested if H_0: $\pi_1-\pi_2=0$ and H_a: $\pi_1-\pi_2\neq 0$. The population proportion of the scores of the pharmacy-mix-related scores of the first 33% (π_1) were assumed to be equal to the scores of the last 33% (π_2). If p-value>α then H_0 was accepted, if p-value<α then H_0 was rejected. If H_0 was to be accepted, independence of the questions was assumed. These hypotheses were tested for all individual cases. For example, for both population proportions of 33% of the questions for case 1, a table was produced to determine the sample proportions.* ***Table 4.4.*** *represents the sample proportions of the first 33% of case 1 and* ***table 4.5.*** *represents the sample proportions of the last 33% of case 1 within 123-method A.*

[14] *Calculated z-value for testing $\pi_1-\pi_2=0$:*

$$z = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}}$$

mix/ranking	1	2	3
product	.53	.22	.25
process	.12	.42	.46
customer	.35	.36	.29

Table 4.4. *Relation of mix and ranking question 1-15 from 123-method A in case 1.*

mix/ranking	1	2	3
product	.41	.34	.25
process	.14	.31	.55
customer	.45	.35	.20

Table 4.5. *Relation of mix and ranking question 31-45 from 123-method A in case 1.*

With the binomial test the score .53 (cell upper left ***table 4.4.****) was compared with the score .41 (cell upper left* ***table 4.5.****). The computed z-value was .83 and was compared with the table of cumulative standardized normal distribution F(z). The according p-value was p= 0.7967. In this example the p-value(.7967)>α(0.05). In addition, within all other tests with respect to the other cases the p-value>α. H_0 was accepted, and independence of the questions was assumed. The pattern in the answers of the population proportion of the first 33% of the answers is not significantly different from the last 33% of the answers. Independence of the first and the last part of the questions seems acceptable. The Friedman test was applied.*[15]

The card-method was processed in exactly the same way as the 123-methods. However, some new problems arose. The 10 cards were divided into the categories 'very important', 'not important' and 'remaining', with a minimum of 3 cards and a maximum of 4 cards per category. Within the category very important the cards were ranked (most to least important). Only the ranked cards were used in the statistical processing. All pharmacy managers ranked the maximum of 4 cards in the category very important. Two problems were identified. The first problem related to the cards not used (= categories not important and remaining). For example, all cards relating to the product mix were selected in the category not important. The four cards in the category very important would get the values 1-4 (most-least important), and we first intended to give the other cards the value 0. In the example, the minimum value of the not used cards was zero. This value would be the highest ranking in a Friedman test. Since the product-mix-related cards were not used in the category very important, the highest ranking was regarded as undesirable. A help variable was introduced to overcome this problem. All cards in the categories not important and remaining were given the value 5, instead of the value 0. However, theoretically, some of the ranking of the cards could cause a problem with equality. For example, very important= product(1)-customer(2)-customer(3)-product(4). The sum of the values for the pharmacy mixes are product mix 15(1+4+5+5), process mix 20(5+5+5+5), and customer 15(2+3+5+5). Were product mix and customer mix equally important? It was accepted as equal, although such rankings were not observed in the pilot study. However, in two cases, equality was present in another form. For example, no cards of the product mix and process mix were used, and resulted in a score of 20 for both mixes. This was accepted as equal also. The second problem related to the number of cards per pharmacy mix. A total of 10 cards were divided into 3 cards for the product mix, 3 cards for the process mix, and 4 cards for the customer mix. As a result, within

[15] *Within this study a test was made with respect to the outcome of the ordinal scale in a Friedman test and was compared by calculating a mean of the values used in the Friedman test, and calculating the mean of other values. If the same data matrix was used, the result of calculating the mean was exactly in correspondence with the result of the Friedman test. In addition, the distance between 1, 2 and 3 does not have a meaning in an ordinal scale. This was tested to be true by calculating the mean for some other values. The values 1, 2 and 3 were transformed into 1, 1.25 and 1.5; 1, 1.5 and 2; 1, 3 and 5 and 1, 5 and 9 respectively. This resulted in exactly the same ranks for all cases. Consequently, the ranks of the Friedman test seemed to be quite stable.*

the Friedman test the range from the best possible score to the lowest possible score varied per mix. The product mix and the process mix ranged from 11(1+2+3+5) to 15(5+5+5). The customer mix ranged from 10(1+2+3+4) to 20 (5+5+5+5). This division of cards resulted in a more detailed scale for the customer mix (10-20) than the scale for the product mix and the process mix (11-15).

With the drawing method a pictogram was produced. All pharmacies were represented by three marks in the triangle of product mix, process mix, and customer mix. The angular starting point, moving along the axe and the diagonal-end point were used in visualizing the rankings.

4.3.3. Action methods

With the questionnaire for ***action***, a total number of 17.204 items was used in the statistical processing. The data were processed with cluster analysis. Selection and collection of data for social items can produce an embarrassment of riches; for example, cumulative scores per pharmacy manager and ratio scores per pharmacy. A method of producing clear comprehensible patterns without losing the contribution from each number is needed. Cluster analysis provides such a method of organizing data (Monti 1975). The methods for the pharmacy manager, the pharmacy personnel and the researcher were all processed with cluster analysis. Some descriptions of cluster analysis are presented.

Cluster analysis is a multi-variate technique, in which each case may describe more variables. Such an analysis divides cases in several groups. The cases within one group should be very alike, and at least not be like cases in another group. The researcher needs to find clusters or groups of cases with common characteristics. In cluster analysis, the number of clusters, the cluster members and the common characteristics have to be determined. It is used in a search for a typology (Slotboom 1987). Kiewiet and Stegwee (1991) and Kiewiet (1996) describe cluster analysis as the dividing of a set of cases into several subsets with the use of similarity measurements. Tryon et al. (1970) describe cluster analysis more broadly, as the general logic, formulated as a procedure, by which we objectively group together entities on the basis of their similarities and differences.

Within this study, an agglomerative hierarchical cluster analysis was used. Within this cluster analysis, two cases are partitioned into one cluster. Consequently, these two cases cannot be partitioned in two different classes later in the calculation. Why a hierarchical cluster analysis? Hierarchical methods give n nested classifications from n clusters of one member each, to one cluster of n members. Nonhierarchical clustering methods are designed to cluster data units into a single classification of n clusters, where n either specified a priori or is determined as part of the clustering method (Anderberg 1973). Within this study the number of clusters was not specified a priori since the number of pharmacy mixes was not fixed a priori but ought to be a result of the clustering of pharmacy-mix-related actions. The number of clusters also is not the main focus of this study, and therefore not directly related to the research question. This study seeks to find clusters stressing the same actions. The number of clusters is of secondary importance. Another question is why there should be an agglomerative hierarchical approach. The agglomerative approach pulls together the entities which are most alike. Other hierarchical clustering methods focus on finding groups which are best separated from each other or most distinctive (Anderberg 1973). Neither separation or distinction is primarily related to the research *question. In this study we intended to select similar cases with respect to pharmacy-mix-related actions. The agglomerative hierarchical cluster analysis was selected as the clustering technique for action.*

Cluster analysis consists of several steps. Anderberg (1973) described nine steps: choice of data units, choice of variables, what to cluster, homogenizing variables, choice of similarity measures, the clustering criterion, algorithms and computer implementation, deciding the number of clusters, interpretation of the results. Romesburg (1984) described several steps, such as obtaining the data matrix, standardizing the data matrix, computing the resemblance matrix, executing the clustering method, rearranging the data and resemblance matrices, and computing the cophenetic correlation coefficient. Other authors described more or less similar steps. In this pilot study, five steps were selected: step 1 the choice of objects and variables; step 2 standardizing the data matrix; step 3 choice of similarity measures; step 4 choice of clustering method; step 5 deciding the number of clusters. All steps will be described below. Some results of the pharmacy manager will presented in order to illustrate the choices made. The remaining questionnaires, for the pharmacy personnel and the researcher, were processed with the same format.

***Step 1** The choice of objects and variables*
Cases of the pilot study were pharmacy managers. Variables were the count variables describing the number of actions related to product, process, customer and remaining.

***Step 2** Standardizing the data matrix*
Per object the scores were calculated. Four categories of answers were used: 'yes', 'no', 'not applicable', and 'missing' in combination with the pharmacy-mix-related actions 'product', 'process', 'customer', 'remaining' and 'total'. Firstly, all scores of the four categories of answers were counted. For example, the overall score for case 1 on the answer 'yes' was 132 out of 372. Secondly, the scores of the four categories of answers were counted per pharmacy-mix-related action. For example, the product score for case 1 on the answer 'yes' was 33 out of 92. The sum total of the scores per category of answers and the sum total of the scores per category of answers per pharmacy mix were calculated separately, and compared. Both totals resulted in 372 questions. This outcome was used to verify this part of the processing. The results were translated from the nominal scale, via a count variable, to ratio scale. This translation was made in order to combine the results of the questions relating to a specific pharmacy mix. The different scores were used as the 'pharmacy-mix'-shape of a case. The shape showed the focus of the case with respect to the three pharmacy mixes. Consequently, the binary scale was not used in the statistical processing of action. The count variable for the cluster analysis was the number of scores 'yes' per pharmacy-mix-related action. Since the number of questions per pharmacy mix were unequal and some questions were 'not applicable' or 'missing', corrections were required. Two corrections were made relating to the maximum score per pharmacy-mix-related action.

The first correction related to the unequal outcome per pharmacy-mix-related action. For example, the uncorrected maximum scores of the pharmacy manager for the actions in product, process, customer, remaining and total were 92, 111, 100, 69 and 372 respectively. To overcome this problem the data matrix was standardized.

Romesburg (1984) argues that standardizing makes attributes contribute more equally to the similarities among objects. For example, if the range of values on the first attribute is much greater than the range of values of the second, then the first attribute will carry more weight in determining the similarities among objects. When this is undesirable in the context of your *research goal, you can compensate for the effect by standardization. Within these methods, all count variables were rearranged on a ratio scale from 0 to 1 to overcome the problem of different ranges of values. The second correction related to the answers 'not applicable' and 'missing'; the categories describe different phenomena. However, both categories were used as one category with respect to their effect on the possible maximum score in the statistical processing. The possible maximum score of the main result of the calculations, the count variable, was affected by these answers. For example, if the answers 'not applicable' and 'missing' were used more often in a certain case, the maximum possible score in that case was lower. Per case a correction was made. For example, a correction for the product mix of case 1 could be made. Normally, by calculating a ratio score of case 1, the product-mix score would be divided by the maximum possible product-mix score. However, in this study the maximum possible score was reduced to the number of questions in the categories 'not applicable' and 'missing'.*

***Step 3** Choice of similarity measures*
A similarity measure is a mathematic formula, and represents how similar the pair of objects are (Romesburg 1984). This study deals with the emphasis of the actions of the pharmacy manager with respect to the three pharmacy mixes. Different similarity measures can express similarity among scores with respect to height, size or shape. The choice of a similarity measure relates to the validity of the study. How well does the measure applied serve the research goal? For example, the question 'What package of business actions is the main focus of the pharmacy manager?', relates to the shape of the ratio scores. Within this part of the study, two measures of validity were used. The primary validity deals with the question how well a cluster analysis achieves its research goal and generates interesting and useful conclusions. The secondary validity deals with certain features we would like the cluster analysis to have (Romesburg 1984). The primary validity will be discussed here. The secondary validity will later be discussed in conjunction with Step 4, the choice of the clustering methods.

The primary validity relates to the similarity measures which are sensitive to the height, size or shape of the scores of the individual pharmacy mixes.[16] Other similarity measures are sensitive to the shape of the scores. The difference in the use of these similarity measures is related to the term 'size displacement'. Romesburg (1984) argued that a size displacement occurs when the data profile of one case is, attribute for attribute, larger or smaller than that of another. Size displacement is a useful concept because we want to measure similarity based on the shapes of the profiles rather than size displacement between them. For example, if the cases are fossil

[16] *Squared euclidian distance* $[(X,Y) = \Sigma(X_i-Y_i)^2]$ *and block* $[(X,Y) = \Sigma|X_i-Y_i|]$ *were the applied similarity measures in the first try-out of the statistical processing. The clustering resulted in a strong positive relation between the **height** of the scores per pharmacy-mix-related ratio. A high score on the product-mix ratio was clustered with a high score on the process-mix ratio, customer-mix ratio and remaining ratio. The use of these similarity measures resulted in a very robust clustering with respect to secondary validity. However the primary validity was low. The cases were divided into a cluster of some (6) young pharmacy managers (low ratio scores), a big cluster (16) of middle-range scoring pharmacy managers (middle ratio scores), and a very small (2) top-segment (high ratio scores). This was no new fact to stichting VNA and SAL Apotheken. In fact, many of the actions of these organizations concentrate on the supporting of young pharmacy managers. No relation to the research question (the emphasis of the actions of the pharmacy manager with respect to the pharmacy mixes) was made. Consequently, the relevancy and utility of the clustering with the similarity measures squared euclidian and block were small.*

bones and the attributes are measurements of the bones, we probably want to ignore size displacement between data profiles: after all, size displacements can arise incidentally because the animals were not all the same age when they died. If it is presumed that the fossil bones were found near an ancient village in the province of Friesland in the Netherlands, we have to be able to distinguish a mammoth from a forest elephant, since, the forest elephant lived 70.000 years earlier than the much bigger mammoth from the tundra (Bos 1995). This information may be of great importance to archaeologists. Whether the bones originated from a young mammoth or from a fully-developed mammoth is not relevant for the used classification; in this case the bones should be classified in the category mammoth. Similarly, the pharmacy manager should be classified in the category product mix, process mix or customer mix. Whether a difference between young pharmacy managers who are relatively new in the profession, or fully-developed pharmacy managers, who are experienced in the profession, is observed, is not of any major importance here. Ideally, the pharmacy manager should be classified in one of the three pharmacy mixes. Romesburg (1984) describes the Pearson correlation coefficient and the cosines coefficient as similarity measures which are sensitive to shape. Both measures are relatively indifferent to size displacement. Linear translations are usually ignored in the Pearson correlation coefficient.[17] The Pearson correlation coefficient was selected as a similarity measure in this study.[18]

Step 4 *Choice of clustering method*
Hierarchical cluster analysis consists of several methods of clustering. A choice must be made among these various methods. The two measures of validity (Romesburg 1984), mentioned before in Step 3, were used for the choice of the clustering methods.

The primary validity, which deals with the question of how well a cluster analysis achieves its research goal and generates interesting and useful conclusions, consisted of a demonstration in the use of the most widely applied clustering methods (Romesburg 1984).

---

[17] *Both additive translations, scores of pharmacy manager A = scores pharmacy manager B + 2, and proportional translations, scores of pharmacy manager A = scores pharmacy manager B x 2, are ignored in the Pearson correlation coefficient. The cosines coefficient ignores proportional translations only.*

[18] *The Pearson correlation describes the correlations between vectors of values.* Z_{Xi} *is the standardized Z score value of X for the ith item, and N is the number of items (Norusis 1993);*

$$Similarity(X,Y) = \frac{\sum_i Z_{X_i} Z_{Y_i}}{N-1}$$

In addition, some of the clustering methods cannot be combined with certain similarity measures.[19] *Within this current study one out of two clustering methods was selected. Both clustering methods were related to average linkage; 'average linkage between groups' and 'average linkage within groups'. Some describe average linkage method as the distance between groups or as the average of the distances between all pairs of individuals in the two groups (Everitt 1980). Sokal and Michener (1958) use the average linkage method as a measure of distance between an individual and a group of individuals, while Lance and Williams (1966) extend it to a measure of distance between groups. In this pilot study, the average linkage* ***between*** *groups method is described as the average of the distances between all pairs of individuals in all groups. In the average linkage* ***within*** *groups method, the distance between groups is described as the average of the distances between all pairs of individuals within each group.*

The first method used is the average linkage between groups or UPGMA (unweighted pair-group method using arithmetic averages) and is a clustering method which defines the distance between two clusters as the average of the distances between all pairs of cases in which one member of the pair is from each cluster. For example, if the cases 1 and 2 of ***figure 4.2.*** *form cluster A and cases 3, 4 and 5 form cluster B, the distance between clusters A and B is taken to be the average of the distances between the following pairs of cases: (1,3)(1,4)(1,5)(2,3)(2,4)(2,5). This differs from the other linkage methods in that it uses information about all pairs of distances, not just the nearest or the furthest. For this reason, it is usually preferred to SLINK and CLINK for cluster analysis (Norusis 1993).*

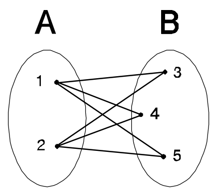

Figure 4.2. *Example of distance between cluster A and B with average linkage between groups.*

The second method used is a variant of UPGMA, the average linkage within groups. It is a clustering method in which a cluster is characterized by the average of all links within it (Anderberg 1973). This method combines clusters so that the average distance between all cases in the resulting cluster is as small as possible. Thus, the distance between two clusters is taken to be the average of the distances between all possible pairs of cases in the resulting cluster (Norusis 1993). For example, if cases 1 and 2 of ***figure 4.3.*** *form cluster A and cases 3, 4 and 5 form cluster B, the distance between clusters A and B is taken to be the average of the distances between the following pairs of cases: (1,2)(3,4)(3,5)(4,5).*

[19] *Within the first try-out of the statistical processing, two of the most widely used clustering methods (Romesburg 1984) average linkage* ***between*** *groups and Ward were applied. The combination of these clustering methods with the similarity measures squared euclidian and block resulted in a strong positive relation between the height of the scores per pharmacy-mix-related ratio mentioned earlier. The primary validity was, however, low. As a result, and in accordance with the arguments mentioned in Step 3, the similarity measure Pearson correlation coefficient was selected as the appropriate measure of association. However, Pearson correlation is a resemblance coefficient. The Ward method cannot be used with any resemblance coefficient. The Ward method was replaced by another average linkage clustering method:* ***within*** *groups. The average-linkage clustering method can be used with resemblance coefficients, and it judges the similarity between pairs of clusters in a manner less extreme than either SLINK (single linkage) or CLINK (complete linkage) (Romesburg 1984).*

The secondary validity, which deals with certain features we would like the clusters to have, consists of a demonstration of robustness and an agreement with expert intuition (Romesburg 1984). In addition, in this study the researcher has made a rough classification. Firstly, Williams (1967) described stability and robustness. In general, we could say that if adding information does not disturb the classification, the clustering is taken to be stable; if removing information does not produce major changes in the classification (for example, removing one or two objects or attributes from the original data matrix) the cluster is taken to be robust. In this study the robustness test was made in such a way that 50% of the attributes (pharmacy-mix-related ratios) was removed. Within seven tests, several combinations of the attributes were made.[20] The effect of attribute change to the clustering was compared. Secondly, Romesburg (1984) quoted several examples of where expert intuition can be used (Hodson et al. 1966, Sneath 1968, Gupta and Huefner 1972, Zajicek et al. 1977). In our project the management of stichting VNA and SAL Apotheken positioned all twelve of the VNA/SAL pharmacies in the pharmacy mix before the pilot study. Five pharmacies were located in the product mix, two pharmacies were located in the process mix, and five pharmacies were located in the customer mix. For example, the pharmacy manager of case 17 was a pharmacy manager in the process mix, according to the management of stichting VNA. Finally, a rough classification was made using a comparison of all pharmacy-mix-related ratio scores. The researcher classified all cases in four clusters. The shape of the scores determined the membership of a case in a cluster. The rough clusters were:

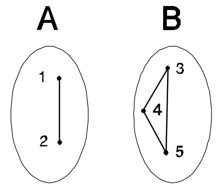

Figure 4.3. *Example of distance between cluster A and B with average linkage within groups.*

cluster 1 (process≈product)>customer
cluster 2 (product≈process)>customer
cluster 3 product>(customer≈process)
cluster 4 product>(process≈customer)

The symbols applied are: > and < for larger and smaller than, ≥ and ≤ for larger and almost equal to and smaller and almost equal to, (mix1≈mix2) two mix scores being almost equal. The results of both clustering methods, average linkage between groups and average linkage within groups, were compared with the criteria of robustness, expert-intuition and rough clustering by the researcher. Within the method of average linkage between groups, 13 cases were replaced after the test for robustness. After the expert-intuition test, nine out of twelve VNA/SAL pharmacies were grouped in according clusters. After the rough-clustering test eighteen cases corresponded with the rough clusters. Within the method average linkage within groups, 12 cases were replaced after the test for robustness. After the expert-intuition test also nine out of twelve VNA/SAL pharmacies were grouped in according clusters. After the rough-clustering test, all cases corresponded with the rough clusters. Consequently, the method average linkage within groups was selected because of the criterion of robustness. The application of this method resulted in four clusters which were in accordance with the results of the rough clusters, seen above.

---

[20] *The original clustering was made with the ratios of product, process, and customer. This result was compared with the clustering of the ratio scores of product-process, product-customer, product-remaining, process-customer, process-remaining, customer-remaining and product-process-customer-remaining.*

***Step 5** Deciding the number of clusters*

*Since the applied hierarchical clustering techniques ultimately reduce the data to a single cluster containing all the entities, the researcher may need to decide at which stage in the analysis he wishes to stop (Everitt 1980). Within this current study the decision about the number of clusters was related to the theoretical expectations of the pilot study. It was expected that three pharmacy mixes would appear. The outcome was not as expected. It was different in such a way that the clustering did not produce pure types as presented in **table 4.1.**, but four clusters with two pairs of clusters similar to that of the rough clusters. The rough clusters had two main features. Firstly, the product-related and process-related actions were nearly equal. Both mix actions were larger than the customer-related actions (cluster A). Secondly, the process-related and customer-related actions were nearly equal. Both mix actions were smaller than the product-related actions (cluster B). Each of these two main clusters were refined. Cluster A was subdivided in two clusters in which process-related actions were a little bit larger than product-related actions (1), vice versa (2). Cluster B was subdivided in two clusters in which customer-related actions were a little bit larger than process-related actions (3), vice versa (4).*

Main cluster A	*cluster 1 (process≈product)>customer*
	cluster 2 (product≈process)>customer
Main cluster B	*cluster 3 product>(customer≈process)*
	cluster 4 product>(process≈customer)

The classification, with the clustering method average linkage within groups and the similarity measure Pearson correlation coefficient, produced clusters corresponding with the rough clusters mentioned before. The rough clustering was used as a criterion for the number of clusters. In addition, the distances in the dendogram, representing the steps in hierarchical clustering method, were chosen as large as possible. Both criteria resulted in four clusters for the questionnaire for the pharmacy manager.

4.4. Results

Methods based on the choices mentioned earlier were applied to the data. The results of this pilot can thus now be presented. The results of the method for *profile* of pharmacy organization will be presented in 4.4.1. The results of the methods for *thought* of the pharmacy manager will be presented in 4.4.2. The results of the methods for *action* of the pharmacy organization will be presented in 4.4.3. The correspondence of the results for *thought* and *action* will be discussed briefly in 4.4.4.

4.4.1. Profile results

The *profile* of pharmacy organization of the selected group was described by the Seniority, Sex and Division of Time of the pharmacy manager, and, in addition, (In)dependence, Cooperation, Location, Part-time/Full-time Personnel, Flow of Prescriptions, Flow of Patients, Turnover and Net Result of the pharmacy organization.

The pharmacy managers were mostly male (71%) and some female (29%). Within the pilot group most of the pharmacy managers had a seniority between 6 and 15 years (62%). Some of the pharmacy managers had an experience of more than 15 years as pharmacy managers (17%), some between 2 and 6 years (13%), and some pharmacy managers were quite new in the field of pharmacy, having less than 2 years experience as pharmacy managers (8%). The time spent on tasks directly relating to the pharmacy organization was more than 40 hours for most of the pharmacy managers (79%) and 30-40 hours for the remaining pharmacy managers (21%). The time spent on tasks indirectly relating to the pharmacy organization was less than 10 hours for most of the pharmacy managers (71%) and 30-40 hours for the remaining pharmacy managers (29%).

Within the pilot group, most of the pharmacy organizations were managed by an independent pharmacy manager (67%). Some of the pharmacy managers were in employment (21%) and, in addition, a small number of pharmacy managers in employment were working towards being independent (12%). Compared with the independence of the pharmacy manager (67%) mentioned before, a smaller number of pharmacies were independent (46%). The remaining pharmacies were participating in a transfer formula (25%), a health-care centre (13%), ownership of several pharmacies (8%), and part of a cooperation of some kind (8%). Most of the pharmacies were not participating in a pharmacy concept (75%). The participating pharmacies were related to various so-called ‘formulas’ (groups of cooperating pharmacists); Meditheek (13%), Kringapotheek (4%) and Extra apotheek (8%). The location of most pharmacies was in the suburban area of a city (42%). Some of the pharmacies were located in the city centre (29%), and in rural areas (25%). Personnel consisted of both part-time and full-time staff. The part-time staff consisted mainly of 5-10 employees (59%); 3-5 employees (29%), 1-2 employees (8%) and 10 or more employees (4%), respectively. The full-time staff consisted mainly of 1-2 employees (33%) and 3-5 employees (33%); 5-10 employees (30%), and 10 or more employees (4%) respectively. Pharmacists and their personnel were both included in the results of Full-Time Equivalence (FTE). The FTE for pharmacists was mainly smaller than 1.5 FTE (79%); 1.5-3 employees (17%). The FTE for other pharmacy personnel was mainly 4-7 FTE (46%); 1-4 employees (33%), 7-9 employees (17%), and 9 or more employees (4%) respectively. The number of prescriptions over 1993 was mainly 50-70.000 prescriptions (29%); 70-90.000 prescriptions (25%), 50.000 prescriptions or less (17%) and 90.000 prescriptions or more (17%) respectively. The number of patients in the whole of 1994 was mainly 8-11.000 patients (54%); 8.000 patients or less (29%) and 11-14.000 prescriptions (17%) respectively. The turnover in the whole of 1993 was mainly 2.5-3.5 million Dutch guilders (38%); 2.5 million or less (25%), 3.5-4.5 million (25%) and 4.5 million or more (4%) respectively. The net profit over 1993 was mainly 250-400.000 Dutch guilders (38%); 100-250.000 (21%), 100.000 or less (17%) and 400.000 or more (17%) respectively.

4.4.2. Thought results

The *thought* of the pharmacy manager was described with the results of the Friedman test by using the data of the 123-methods A and B and the card method. In addition, a pictogram described the results of the drawing method. For all combinations of *profile* and *thought* used, the time spent to fill in the questionnaire was indicated by the respondents. The method profile, combined with 123-method A and drawing method took an average of 29 minutes, the method profile combined with 123-method A and card method took an average of 35 minutes. The method profile combined with 123-method B and drawing method took an average of 51 minutes, the method profile combined with 123-method B and card method took an average of 43 minutes.

The **123-method A** consisted of 28 main questions in which the relation between the question and the pharmacy mix was described. All main questions consisted of three sub-questions relating to a specific pharmacy mix. In the Friedman test, a mean ranking was calculated with an according p-value. The results of this method, presented in **table 4.6.**, consisted of three categories of priorities. The first category related to the ranking customer-product-process (case 3, 16 and 17). The second category related to the ranking product-customer-process (case 6, 9, 11, 14, 18, 19 and 23). The third category related to the ranking product-process-customer (case 1). In the situation of the ranking being equal, for example, as observed in the ranking (1=2) 3, that specific case was allocated to the most frequent ranking within that method, ultimately.[21] In addition, a p-value was calculated. The level of significance, for example, α=0.05, explains the probability that the given rankings, most likely, were not coincidental (Slotboom 1987). The calculated p-values of 123-method A were between .63 and 0.

<table>
<tr><th>ranking
method</th><th>customer product process</th><th>customer process product</th><th>product customer process</th><th>product process customer</th><th>process product customer</th><th>process customer product</th></tr>
<tr><td>123-method A</td><td>case 3, 16, 17</td><td></td><td>case 6, 9, 11, 14, 18, 19, 23</td><td>case 1</td><td></td><td></td></tr>
<tr><td>123-method B</td><td>case 2, 12, 13, 15, 21, 22</td><td></td><td>case 5, 7, 8, 10, 20, 24</td><td></td><td></td><td></td></tr>
<tr><td>card method</td><td>case 5, 7, 8, 12, 16, 17, 19, 23</td><td>case 2, 14</td><td>case 10, 18</td><td></td><td></td><td></td></tr>
<tr><td>drawing method</td><td colspan="2">case 3, 4, 9, 13, 20, 21, 22</td><td colspan="2">case 1, 6, 11, 24</td><td colspan="2">case 15</td></tr>
</table>

Table 4.6. Relation of ranking and *thought* methods.

[21] Within 123-method A, case 6 was allocated to the rank product-customer-process.

The **123-method B** consisted of 45 main questions in which the relation between the question and the pharmacy mix was *not* described. All main questions consisted of three sub-questions relating to a specific pharmacy mix. The results of this method, presented in **table 4.6.**, consisted of two categories of priorities. The first category related to the ranking customer-product-process (case 2, 12, 13, 15, 21 and 22). The second category related to the ranking product-customer-process (case 5, 7, 8, 10, 20 and 24). The calculated p-values of 123-method B were between .05 and 0.

The **card method** consisted of two sets with a total of 10 cards each. Every card related to a specific pharmacy mix. The results of this method, presented in **table 4.6.**, consisted of three categories of priorities. The first category related to the ranking customer-product-process (case 5, 7, 8, 12, 16, 17, 19 and 23). The second category related to the ranking product-customer-process (case 10 and 18). The third category related to the ranking customer-process-product (case 2 and 14). In addition, three cases were allocated.[22] The calculated p-values within the card method were between .22 and .13. The complete ranking of the card method was for 42% (5 cases) corresponding to the results of the 123-methods (**table 4.6.**).

The **drawing method** consisted of three questions related to three positions in the picture of the triangle. With the first question the pharmacy manager positioned the pharmacy in the angular points of the triangle. With the second question the pharmacy manager positioned the pharmacy alongside the axes of the triangle, or, if there was no movement, in the angular points. With the third question the pharmacy manager positioned the pharmacy alongside the diagonal lines, or, if there was no movement, in the angular points or alongside the axes. Within **figure 4.4.** the starting point of the arrow relates to the first question, the crossing of the axes relates to the second question, and the endpoint of the arrow relates to the third question. **Figure 4.4.** shows three positions per pharmacy, and two positions for pharmacy 3 and 21 (after question 2 there was no movement according to the pharmacy manager). **Figure 4.4.** also shows that most of the pharmacies start in the

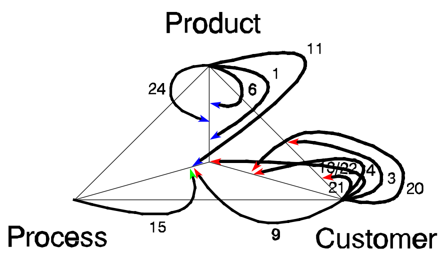

Figure 4.4. *The results of the drawing-method priorities in dynamical perspective: the angular starting point, crossing the axe*, the *diagonal-line end point.*

[22] Within the card method, case 8, 17 and 19 were allocated to the rank customer-product-process.

customer mix, move alongside the axe with the product mix and end by moving over in the direction of the process mix. The results of the drawing method are also described in **table 4.6.** Since the method is hardly comparable with the other methods, only the first ranking is presented. A comparison of the 'first-ranking only' showed that, within the drawing method, 67% of the cases (8 cases) corresponded with the outcome of the 123-methods. The results of the complete ranking of the drawing method, although hardly comparable, corresponded for only 8% (1 case) to the results of the 123-methods.

4.4.3. Action results

The *action* of the pharmacy manager was described by using the results of the questionnaire for the functions pharmacy manager, pharmacy personnel (second pharmacist and assistant pharmacist) and researcher. For all questionnaires used for *action*, the time spent to fill in the questionnaire was also indicated by the respondents. The method for the pharmacy manager took an average of 47 minutes, the method for the second pharmacist took an average of 45 minutes, and the method for the assistant pharmacists took an average of 38 minutes. The time spent by the researcher took a mean of 96 minutes per interview. The results of all methods were produced using a cluster analysis. It should be remembered that we decided to use the similarity measure Pearson correlation coefficient and the method average linkage within groups in the cluster analysis. The results of cluster analysis were presented visually in a dendogram which represents the steps in a hierarchical clustering solution. The dendogram identifies the clusters being combined and the values of the coefficients at each step. In this method, the distances in the presented dendograms are rescaled to numbers between 0 and 25. The ratio of the distances between steps is preserved, but the scale displayed at the top of the figure does not correspond to actual distance values (Norusis 1993). The rough clusters, mentioned earlier, were used as a criterion for the number of clusters. In addition, the distances in the dendogram were chosen to be as large as possible, and, are equal for the three methods.

With the questionnaire for the **pharmacy manager**, as mentioned earlier, the main results of the *rough clustering* were twofold. Firstly, the product-related and process-related actions were nearly equal. Both mix actions were larger than the customer-related actions. Secondly, the process-related and customer-related actions were nearly equal. Both mix actions were smaller than the product-related actions. Each of these two main clusters were refined. The first main cluster was subdivided into two clusters in which process-related actions were slightly larger than product-related actions (cluster 1) and vice versa (cluster 2). The second main cluster was also subdivided into two clusters in which customer-related actions were slightly larger than process-related actions (cluster 3) and vice versa (cluster 4).

The final result of the *empirical clustering* for the pharmacy manager was also four clusters (**figure 4.5.**) ranked thus: *cluster 1* (process≈product)>customer
cluster 2 (product≈process)>customer
cluster 3 product>(customer≈process)
cluster 4 product>(process≈customer)

We also chose the distances in the dendogram to be as large as possible. In this method the rescaled distance cluster combine was chosen to be just under 5. This is marked by a vertical dotted line in **figure 4.5.** resulting in four clusters.

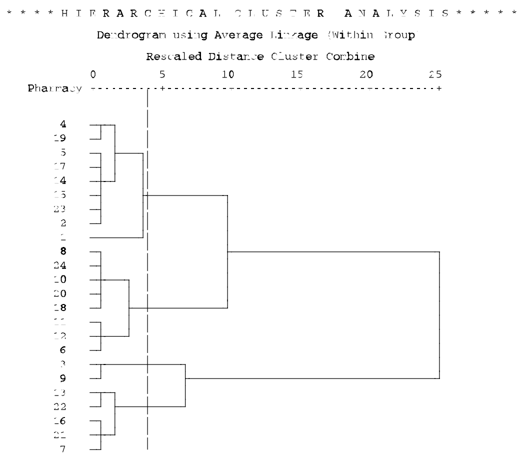

Figure 4.5. The dendogram representing four clusters from the questionnaire for the pharmacy manager.

With the questionnaire for the **pharmacy personnel** the results of the *rough clustering* were also mainly twofold. Firstly, product-related actions were larger than process-related and customer-related actions. Secondly, process-related actions were larger than product-related and customer-related actions. Each of these two main clusters were refined. The first main cluster was subdivided into a cluster in which product-related actions were larger than process-related actions which in turn were larger than customer-related actions; all to a different extent (cluster 1), and, in addition, was subdivided into a cluster in which product-related actions were larger than customer-related actions which in turn were larger than process-related actions (cluster 2); all to a different extent also. The second main cluster was subdivided into a cluster in which process-related actions were larger than product-related actions which in turn were larger than customer-related actions (cluster 3), and, was subdivided into a cluster in which process-related actions were larger than the customer-related actions which in turn were larger than product-related actions (cluster 4), and, finally, was subdivided into a cluster in which process-related actions were slightly larger than the product-related actions, both actions being larger than the customer-related actions (cluster 5).

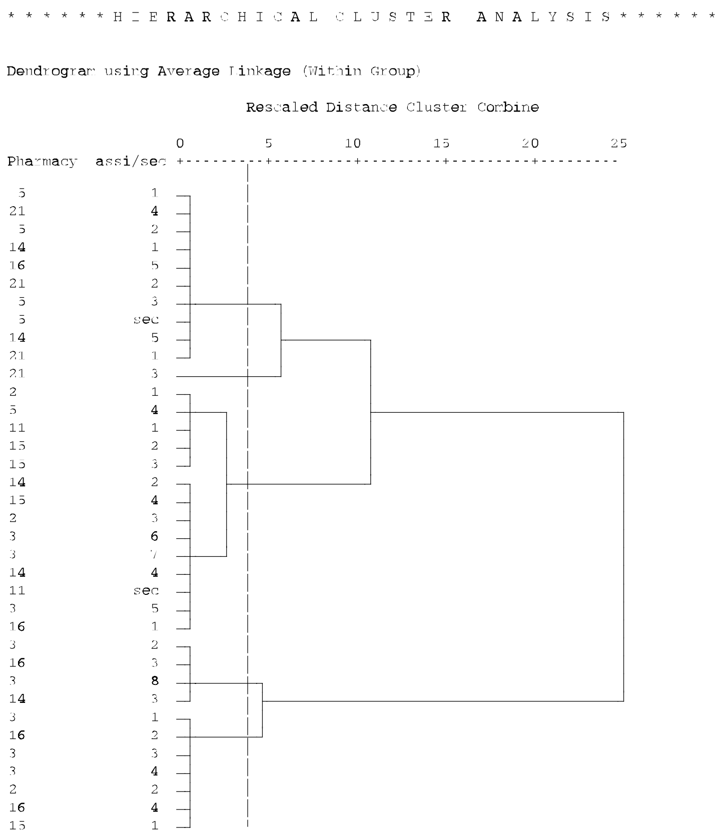

Figure 4.6. The dendogram representing five clusters from the questionnaire for the pharmacy personnel.

The final result of the *empirical clustering* for the pharmacy personnel was also five clusters (**figure 4.6.**) ranked thus:

cluster 1 product≥process≥customer
cluster 2 product≥customer≥process
cluster 3 process>product>customer
cluster 4 process>customer>product
cluster 5 (process≈product)>customer

Three out of 36 cases were placed in a cluster which did not correspond to the rough clusters. Case 15.3 was positioned in cluster 5 with cluster analysis and in cluster 1 with the rough clustering. Case 3.8. and 14.3. were positioned in cluster 2 with cluster analysis and in cluster 1 with the rough clustering. In addition, the distances in the dendogram were to be chosen exactly the same as in the clustering of the pharmacy manager; the rescaled distance cluster combine was chosen to be just under 5, and also marked by a vertical dotted line in **figure 4.6.** resulting in five clusters. The clusters of the pharmacy personnel were compared with the ‘natural cluster’. A natural cluster would be observed whenever the pharmacy personnel of one pharmacy was in the same calculated cluster. However, the respondents of the same pharmacy were never observed in one calculated cluster. Also, the complete clustering of the pharmacy personnel corresponded for only 13% to the results of the pharmacy manager.

With the questionnaire for the **researcher**, the results of the *rough clustering* were also mainly twofold. Firstly, product-related actions were larger than process-related and customer-related actions (cluster 1). Secondly, process-related actions were larger than product-related and customer-related actions. Only the second main cluster was refined. The second main clustering was subdivided into three clusters in which process-related actions were slightly larger than product-related actions (cluster 2), product-related actions were slightly larger than customer-related actions (cluster 3), and, finally, process-related actions were larger than product-related actions which in turn were larger than customer-related actions (cluster 4).

The final result of the *empirical clustering* for the researcher was four clusters (**figure 4.7.**) ranked thus:

cluster 1 (product≈process)>customer
cluster 2 (process≈product)>customer
cluster 3 process>(product≈customer)
cluster 4 process>product>customer

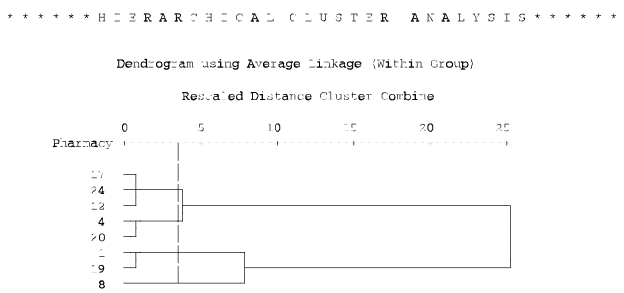

Figure 4.7. The dendogram representing four clusters from the questionnaire for the researcher.

In addition, the distances in the dendogram were chosen to be exactly the same as in the clustering of the pharmacy manager and the pharmacy personnel; the rescaled distance cluster combine was chosen to be just under 5, and again marked by a vertical dotted line in **figure 4.7.** resulting in four clusters. The complete clustering for the researcher was for 63% corresponding to the results of the pharmacy manager.

A summary of the former results of the methods for *action*, as decided by the pharmacy manager, pharmacy personnel, and researcher is presented in **table 4.7.** The refined rankings were transformed into four categories in order to enable a comparison between the three methods. The results are summarized in four main rankings of clusters: product-process-customer, product-customer-process, process-product-customer and process-customer-product.

ranking / method	***product process customer***	***product customer process***	***process product customer***	***process customer product***
pharmacy manager	case 6, 7, 8, 10, 11, 12, 13, 16, 18, 20, 21, 22, 24	case 3, 9	case 1, 2, 4, 5, 14, 15, 17, 19, 23	
pharmacy personnel	case 2.3, 3.5, 3.6, 3.7, 14.2, 14.4, 15.4, 16.1, sec.11	case 2.2, 3.1, 3.2, 3.3, 3.4, 3.8, 14.3, 15.1, 16.2, 16.3, 16.4	case 2.1, 5.1, 5.2, 5.3, 5.4, sec.5, 11.1, 14.1, 14.5, 15.2, 15.3, 16.5, 21.1, 21.2, 21.4	case 21.3
researcher	case 12, 17, 24		case 1, 4, 8, 19, 20	
total cases	**25 cases**	**13 cases**	**29 cases**	**1 case**

Table 4.7. Relation of ranking and *action* methods.

Table 4.7. shows that most cases were observed in the ranking process-product-customer (29 cases), followed immediately by the ranking product-process-customer (25 cases). One can see that all three methods produce both rankings. The third ranking was product-customer-process (13 cases) and was only observed in the method for the pharmacy manager and the pharmacy personnel. The fourth ranking was process-customer-product (1 case) and was only observed in the method for the pharmacy personnel.

4.4.4. Correspondence of results

There is some variety in the results of the methods applied. How do *thought* and *action* correspond with respect to the methods applied? **Table 4.8.** shows the overall results of the four methods for *thought* and three methods for *action*. The results of the drawing method were only included to compare the first ranking. It was argued above that the complete ranking of the drawing method was hardly comparable to the other methods.

The second column of **table 4.8.** product-process-customer, shows that the number of cases for *thought* (1 case) contrasted with the number of cases for *action* (25 cases). Within this ranking, this case 1 was only observed in *thought*. The third column, product-customer-process, shows that the number of cases for *thought* (15 cases) corresponded to the number of cases for *action* (13 cases). However, within this ranking, only two cases were observed in both *thought* and *action*, case number 9 and 14 within 123-method A; case 9 was also observed in the method for the pharmacy manager, and case 14.3. in the method for the pharmacy personnel. The first-ranking-only result of the drawing method shows four cases with 'product' in the first ranking, three of which (cases 6, 11 and 24) were in correspondence with *action*.

ranking **method**	***product process customer***	***product customer process***	***process product customer***	***process customer product***	***customer product process***	***customer process product***
123-method A	case 1	case 6, 9, 11, 14, 18, 19, 23			case 3, 16, 17	
123-method B		case 5, 7, 8, 10, 20, 24			case 2, 12, 13, 15, 21, 22	
card method		case 10, 18			case 5, 7, 8, 12, 16, 17, 19, 23	case 2, 14
drawing method	case 1, 6, 11, 24		case 15		case 3, 4, 9, 13, 20, 21, 22	
total completely ranked cases *thought*	**1**	**15**	**0**	**0**	**17**	**2**
pharmacy manager	case 6, 7, 8, 10, 11, 12, 13, 16, 18, 20, 21, 22, 24	case 3, 9	case 1, 2, 4, 5, 14, 15, 17, 19, 23			
pharmacy personnel	case 2.3, 3.5, 3.6, 3.7, 14.2, 14.4, 15.4, 16.1, sec.11	case 2.2, 3.1, 3.2, 3.3, 3.4, 3.8, 14.3, 15.1, 16.2, 16.3, 16.4	case 2.1, 5.1, 5.2, 5.3, 5.4, sec.5, 11.1, 14.1, 14.5, 15.2, 15.3, 16.5, 21.1, 21.2, 21.4	case 21.3		
researcher	case 12, 17, 24		case 1, 4, 8, 19, 20			
total cases *action*	**25**	**13**	**29**	**1**	**0**	**0**

Table 4.8. Relation of main rankings in *thought* and *action* methods.

The fourth column, process-product-customer, shows that the number of cases for *thought* (0 cases) did not correspond at all with the number of cases for *action* (29 cases). Within this ranking, none of the cases were observed in both *thought* and *action*. The fifth column, process-customer-product, shows that the number of cases for *thought* (0 cases) did not correspond with the number of cases for *action* (1 case). Within this ranking, the case observed in *action* was not observed in *thought*. The first-ranking-only result of the drawing method shows one case with 'process' in the first ranking; this case (case 15) was in correspondence with *action*. The sixth column, customer-product-process, shows that the number of cases for *thought* (17 cases) did not correspond at all with the number of cases for *action* (0 cases). Within this ranking, none of the cases were observed in both *thought* and *action*. The seventh column, customer-process-product, shows that the number of cases for *thought* (2 cases) did not correspond at all with the number of cases for *action* (0 cases). Within this ranking, none of the cases were observed in both *thought* and *action*. The first-ranking-only result of the drawing method shows seven cases with 'customer' in the first ranking, none in correspondence with *action*.

In addition, VNA/SAL pharmacies (case numbers 13-24) were compared with the additional pharmacies (case numbers 1-12). Related to *thought* all VNA/SAL pharmacies were observed in two ranking categories within the 123-methods. Firstly, 6 cases with the ranking product-customer-process (cases 14, 18, 19, 23, 20 and 24) were observed. Secondly, 6 cases with the ranking customer-product-process (cases 16, 17, 13, 15, 21 and 22) were observed. The additional pharmacies were observed in three ranking categories within the 123-methods. Firstly, 1 case with the ranking product-process-customer (case 1) was observed. Secondly, 7 cases with the ranking product-customer-process (cases 6, 9, 11, 5, 7, 8 and 10) were observed. Thirdly, 3 cases with the ranking customer-product-process (cases 3, 2 and 12) were observed. One case was missing (case 4). *Non*-correspondence was observed for most cases (22). Correspondence of *thought* and *action* was observed for one of the VNA/SAL pharmacies (case 14) and for one of the additional pharmacies (case 9).

If all results are compared it can be concluded that the complete ranking (for example, product-process-customer) for *thought* and *action* corresponded in only two cases (cases 9 and 14). The ranking of the first two rankings (for example, product-process) for *thought* and *action* also corresponded in the same two cases. The use of only the first ranking (for example, product) for *thought* and *action* corresponded in 11 cases. This correspondence was observed in the ranking of 'product' in the cases 6-11, 14, 18, 20 and 24, and in the ranking of 'process' in case 15. The

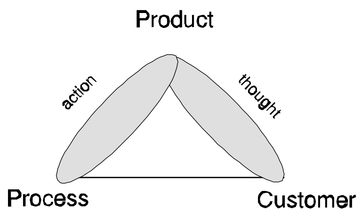

Figure 4.8. The triangle representing the main result of correspondence between *thought* and *action*.

results of the first-ranking-only analysis showed that the main focus of *thought* was related to actions directed at customer and product. In the method for *action* the emphasis of the actions was mainly directed at product and process (**figure 4.8.**). In addition, results of VNA/SAL pharmacies were compared with the results of additional pharmacies. As far as the complete ranking is concerned within all methods, *non*-correspondence was observed for most cases (22). Complete correspondence of *thought* and *action* was observed for one of the VNA/SAL pharmacies (case 14) and for one of the additional pharmacies (case 9).

4.5. Evaluation of methods

The results showed that the outcome of the different methods for *thought* varied. Firstly, the complete ranking of the card method corresponded for 50% with the results of the 123-methods. The results of the drawing method were, although very informative, in contradiction with the results of the 123-methods. The complete ranking of the drawing method, although hardly comparable, only corresponded by 8% to the results of the 123-methods. Secondly, another comparison showed that the p-values of the 123-method A were between .63 and 0; the p-values of the 123-method B were between .05 and 0; and the p-values of the card method were between .22 and .13. Within the drawing method, no p-value was calculated. Thirdly, within the 123-method A, the mode of the p-values was .01 (5 cases). Within the 123-method B, the mode was also .01 (8 cases), and within the card method the mode was .13 (7 cases). Fourthly, for 123-method A and the card method, equal rankings were observed. Within 123-method A, 8% was allocated and within the card method 25% was allocated (you are reminded that allocations were made to the most frequent ranking within that method). Fifthly, interpretation problems in the card method, with respect to the help variable and the different ranges of possible scores between product and process versus customer, did not arise in the other methods. Within 123-method B no cases were allocated, none of the mentioned interpretation problems was observed, and the best p-values (when compared to a frequently applied norm of α=.05) were observed. It was assumed for the moment that the 123-method B would be selected as the instrument for the large sample in the next part of *phase one*. However, after the test for reliability, this assumption would be changed in favour of 123-method A.

The results showed that the outcome of the different methods for *action* also varied. The complete clustering of the method for pharmacy personnel, using the results per respondent, corresponded only for 13% to the results of the pharmacy manager. However, the complete clustering for the researcher corresponded by 63% to the results of the method for the pharmacy manager. The clusters of the pharmacy personnel were compared with the 'natural clusters', meaning that the pharmacy personnel of one pharmacy was in the same natural cluster. The calculated clusters did not correspond with the natural clusters for any pharmacy: the results of the questionnaire for the pharmacy personnel per pharmacy was spread over at least two clusters. This could cause problems with the interpretation of the results if this method was applied to the large sample. In addition, within the method for the pharmacy personnel some cases (8%) were placed in a cluster which did not correspond with the rough clusters. None of these problems of interpretation of data were observed within the method for pharmacy manager and the method for the researcher. In addition, all cases of the method for the pharmacy manager and for the researcher corresponded to the rough clusters. It should however be noted that some of these problems in the method for pharmacy personnel could be solved, for example, by calculating a mean for all respondents per pharmacy, instead of all individuals separately. If this were the case, correspondence between the results of the pharmacy personnel and the pharmacy manager would have improved to 50%. You are however reminded that, the method for the pharmacy personnel and the researcher were only introduced in order to avoid the problem of socially desirable answers; the pharmacy manager selecting questions in accordance with correspondence. In this situation, if correspondence between *thought* and *action* were observed, it would still be possible to test the plausibility of this outcome with the outcome of the two additional methods. Let us, therefore, analyze the correspondence between *thought* and *action* and determine if the method for the pharmacy personnel or the researcher should be used in the next part of *phase two*, the large sample.

In order to give a complete picture of the differences between *thought* and *action*, the results for both 123-methods were used in a comparison with the results of all methods for *action*. Firstly, if the results of the 123-methods are compared with the results of the pharmacy manager non-correspondence was observed for 92% of the pharmacies, correspondence for 4% of the pharmacies, and 4% was missing. Secondly, if the results of the 123-methods are compared with the results of the pharmacy personnel, non-correspondence was observed for 38% of the pharmacies, and 62% of the pharmacies were not comparable with the 123-methods since various rankings existed per pharmacy. Thirdly, if the results of the 123-methods were compared with the results of the researcher non-correspondence was observed for 100% of the pharmacies. If all results were compared, the ‘complete’ ranking (for example, product-process-customer) for *thought* and *action* corresponded for only two cases; the ‘double’ ranking (for example, product-process) for the first two rankings for *thought* and *action* also corresponded for the same two cases; and the ‘single’ ranking (for example, product) for the first ranking only for *thought* and *action* corresponded for 11 cases. The results of the ‘single’-ranking analysis showed that the main focus of *thought* was related to actions directed at customer and product. In the method for *action*, the emphasis of the actions was mainly directed at product and process.

The method for the pharmacy manager was the main instrument for the measurement of *action*. It was anticipated that the pharmacy manager might select questions in accordance with correspondence. In order to avoid this problem, other methods for the pharmacy personnel and for the researcher, were introduced. But if *non*-correspondence were observed, it would be assumed that no further comparisons were required. If the results of the 123-method B are compared with, for example, the results of the pharmacy manager, non-correspondence was observed for 100% of the pharmacies. This means that the rankings of the 123-method B; for example for case 2, customer-product-process, did not correspond with the outcome of the method for the pharmacy manager. The method for the pharmacy manager resulted in a ranking of process-product-customer. *Non*-correspondence between the outcome of 123-method B and the method for the pharmacy manager was observed. In addition, no interpretation problems with respect to the 'natural clusters' existed. For the moment, the method for the pharmacy manager was selected. As a result, it was assumed *for this moment*, that the 123-method B could be selected to visualize *thought* and the method for the pharmacy manager was selected to visualize *action*. However, it will be shown below that following the analysis for reliability, the other method for *thought*, the 123-method A, was preferable.

To make sure that the study, if repeated, would produce the same results, reliability was tested. A method is reliable if the number of accidental errors is minimized (Swanborn 1987). Would the test yield the same score for an individual if two or more measurements are made (Kolstoe 1973)? For example, measures of reliability relate to the internal consistency of a test (e.g. Cronbach's alpha); to splitting the test in two parts and comparing the correlation between the two parts (e.g. split-half); or to models in which extra information about the items is present, for example, in which the items are assumed to have the same means and variances for true (unobservable) scores (e.g. strictly parallel) (Norusis 1993). The split-half model has the disadvantage that the results depend on the allocation of items to the halves. Cronbach's alpha can be applied to all items and no extra information about the items need be present, which was necessary for the strictly-parallel model. Therefore Cronbach's alpha was selected as the reliability test.

The reliability of this study is expressed in Cronbach's alpha for the selected methods for thought and action of pharmacy managers. Cronbach's alpha is based on the internal consistency of the test. Within this method reliability is based on the average correlation of items within a test. Alpha can be viewed as the correlation between this scale and all possible scales that contain the same number of items, which could be constructed from a hypothetical universe of items that measure the characteristic of interest. For example, the set of selected questions for the product mix is a sample from a universe of many possible items. It is assumed that the items on a scale are positively correlated with each other, because they are measuring, to a certain extent, a common entity. Cronbach's alpha tells us how much correlation we expect between our scale of the product mix and all other possible scales measuring the product mix. Since alpha can be interpreted as a correlation coefficient, it ranges in value from 0 to 1 (Norusis 1993).

Firstly, the reliability test was made with respect to the previously selected method for ***thought*** *of pharmacy managers; the 123-method B. All items which related to a certain pharmacy mix* *were tested jointly and separately. The result of this test was to be used to delete questions which reduced the value of Cronbach's alpha. However, the reliability test within thought showed a very surprising result. The result of Cronbach's alpha for all questions of the 123-method B was -1.78. This unexpected result did not range from 0 to 1. The overall alpha of the 123-method B was extremely low. Many low correlation values and negative correlation values were observed, indicating that the applied scale in this method was not very consistent. As a result, the expected correlation between the applied scale of all three pharmacy mixes in the 123-method B and all other possible scales measuring all three pharmacy mixes was extremely low. Consequently, it was questionable if all three pharmacy mixes could be viewed as one test. They do not seem to be related in any way. One of the main causes of this negative result could be due to the application of 'different scales' with respect to product, process and customer, all used in one test. The 'different scales' ought to be analyzed separately. Within 123-method B, all mix-related questions were analyzed separately. As a result the alpha for the product mix, the process mix and the customer mix was .85, .35 and .80 respectively. The alpha value of the process mix of 123-method B was low, indicating that the applied scale was not very reliable. The alpha value of the product mix and the customer mix were quite good. We asked ourselves how we could improve the value of alpha? Consequently, deletion of negatively correlated questions could increase the value of alpha. There was however a problem with the deletion of questions. All three mix-related sub-questions were coupled within one block of main questions. For example, deletion of questions with respect to the 'low-correlating' process mix would involve the 'better-correlating' questions of the product mix and the customer mix also. If all negative correlations were deleted, only 12 out of 45 questions would remain in the questionnaire. The alpha would thus decrease to -4.73 if the negative-correlating questions of the process mix were deleted. The decrease of alpha was mainly due to the deletion of the coupled 'better-correlating' questions of the product mix and the customer mix. As a result, the alpha for the process mix would increase to .56, at the expense of a decrease of the product mix and the customer mix to .77 and .74, respectively. These results showed that the 123-method B did not seem to be a good alternative with respect to reliability.*

Consequently, the 123-method A was introduced and analyzed again. All mix-related questions were analyzed separately. As a result the alpha for the product mix, the process mix and the customer mix was .91, .96 and .94 respectively. The alpha value of the pharmacy mixes, measured separately, were high in 123-method A, indicating that the applied scale was reliable. If the negative correlations were deleted 24 out of 28 questions would remain in the questionnaire. The corrected alpha of all questions would be .98. After the correction, the alpha for the product mix, the process mix and the customer mix would be .93, .97 and .96 respectively. Note that the values of the pharmacy mixes of 123-method A were large, indicating that the applied scale was more reliable after the correction.

Reliability of 123-method A was much better than reliability of 123-method B. But the results with respect to the applied Friedman test showed that the calculated p-value was lower within 123-method A, than within 123-method B. Moreover, cases with equal rankings were observed; two cases were allocated (to the most frequently observed ranking) within 123-method A, and no cases were allocated within 123-method B. However, the results of the corrected 123-methods were compared also. The Friedman test was made with the corrected 123-method A and 123-method B. As a result, the performance of the 123-method B declined, and the performance of the 123-method A improved.

The calculated p-value of the 123-method A was between .63 and 0 within the uncorrected questionnaire, and improved to a range between .42 and 0 when the questionnaire was corrected *with the results of the reliability test. In addition, within 123-method A p<.01 was observed in 7 cases within the uncorrected questionnaire, and remained stable at a value of p<.01 for 7 cases within the corrected questionnaire. Cases with equal rankings were observed within the 123-method A. Two cases were allocated (to the most frequently observed ranking) within 123-method A, within the uncorrected questionnaire, and remained stable with two allocated cases within the corrected questionnaire. The calculated p-value of the 123-method B was between .05 and 0 within the uncorrected questionnaire, and declined to a range between .78 and .01 after the questionnaire was corrected using the results of the reliability test. In addition, within 123-method B p<.01 was observed in 10 cases within the uncorrected questionnaire, but was observed at 6 cases within the corrected questionnaire. Cases with equal rankings were not observed within the uncorrected 123-method B. However, within the corrected 123-method B, one case was allocated (to the most frequently observed ranking) within 123-method B. Finally, if the results of the 123-method A are compared to the results of the pharmacy manager, non-correspondence was observed for 92% (11) of the pharmacies involved. Correspondence of the results was observed for only one case (case 9). This means that the rankings of the 123-method A did not correspond in general to the outcome of the method for pharmacy manager. Non-correspondence between the outcome of 123-method A and the method for the pharmacy manager was observed. With these new results, the 123-method A and the 123-method B were compared. In general, the performance of the 123-method A improved, and the performance of the 123-method B declined. The corrected results showed that the range in the p-value was smaller within the 123-method A (.42-0) than within the 123-method B (.78-.01) for the corrected questionnaire. In addition, the p-value of p<.01 was observed more frequently within 123-method A (7 cases) than within 123-method B (6 cases). In both 123-method A and 123-method B equal rankings were observed. Finally, the results for the pharmacy mixes product, process and customer showed a better performance with respect to Cronbach's alpha within 123-method A (.93, .97 and .96 respectively) than within 123-method B (.77, .56 and .74 respectively). The 123-method A was selected as the instrument to measure thought for the large sample in the next part of phase one.*

Secondly, the reliability test was also made with respect to the method for ***action*** *of pharmacy managers. All items relating to a certain pharmacy mix were tested together, separately, and in combination with the remaining questions. In this study, the values of the inter-item-correlation coefficients per question for each pharmacy mix were quite small with the corrected questionnaire. Norusis (1993) noted that large reliability coefficients can be produced, even when the average inter-item correlation is small, if the number of items on the scale is large enough. Within this study the total number of items was large. The result of Cronbach's alpha for the product mix, the process mix and the customer mix was .86, .75 and .86 respectively. The result of this test was used to delete and transpose questions for the purpose of statistical processing, which would raise the value of Cronbach's alpha. Five questions were transposed from the category 'remaining' to the product mix, and 25 product-related questions were deleted in this mix. Three questions were transposed from the category 'remaining' to the process mix, and 36 process-related questions were deleted in this mix. Four questions were transposed from the category 'remaining' to the customer mix, and 26 customer-related questions were deleted in this mix. As a result a total of 87 pharmacy-mix-related questions were deleted and 12 remaining questions were transposed for the complete questionnaire. In addition, other criteria were used in order to delete or replace the questions. Firstly, if the outcome was exactly the same for all cases (all answers 'yes' or all answers 'no') the result would be hardly usable. All outcome in these categories were redefined as remaining questions, since no relation to any of the pharmacy mixes was possible. Secondly, some combinations of questions, used to check the correctness of the outcome ('yes' here combined with 'no' there), were deleted. The expected result was not obtained. Only one pair of such questions was used in the statistical processing of pharmacy* *manager, second pharmacist and assistant pharmacist. Thirdly, the questionnaires consisted of binary questions (true/false). With this method, a nominal scale was used, and, in addition, the questions were rescaled via a count variable to ratio scale. Three of the questions were deleted since it was not possible to translate these questions from a nominal scale, via a count variable, to a ratio scale. Fourthly, five questions were added following the suggestions of the pharmacy managers. The final result of Cronbach's alpha for the product mix, the process mix and the customer mix was .91, .88 and .91 respectively. Note that the values of the alpha for action were also large, indicating that the applied scale was quite reliable after the correction. The method for the pharmacy manager was selected as the instrument to measure action of the large sample in the next part of phase one.*

All methods applied were evaluated. The 123-method A and the method for the pharmacy manager showed the best performance relatively. In addition, *non*-correspondence between the outcome of 123-method A and the method for the pharmacy manager was observed. An initial difference between *thought* and *action* was determined. The 123-method A was selected to measure *thought* and the method for the pharmacy manager was selected to measure *action*. Both methods will be used in the questionnaire for the large sample in the next part of *phase one*.

4.6. Conclusion

The pilot study was made at 24 pharmacies in the Netherlands. The number of pharmacies included in this pilot was small. With this study it is not possible to generalize to the population of Dutch pharmacies. For that purpose, a large sample will be used in the next part of *phase one* of this thesis. It should be remembered that the main research question for this present chapter, mentioned in **chapter 2**, was: 'What methods can be best applied in a survey?'. The aim of this pilot study was to compare the results of several methods. Moreover, the pilot study sought to find out which of the used methods could be applied best to identify correspondence of *thought* and *action* of pharmacy managers. Additionally, it was examined whether the identified pharmacy mixes were internally consistent; a sketch of correspondence of *thought* and *action* of pharmacy managers was made.

First, we determined what methods could be best applied to identify *thought* of the pharmacy manager in the pharmacy mix for a large sample. The rankings of the drawing method (although hardly comparable), and the card method did not correspond with the 123-methods completely. After the correction of the questionnaire with the results of the reliability test, the calculated p-value was better in the 123-method-A, than the 123-method-B and the card method; in the drawing method no p-value was calculated. Moreover, the highest mode of the p-value was observed at the 123-method A. However, it should be noted that with a frequently applied norm of α=.05, some of the calculated rankings would remain *in*significant. Within 123-method A, 123-method B and the card method, equal rankings were observed; cases were allocated to the most frequently observed rankings. Within the drawing method, no cases were allocated. The card method produced interpretation problems with respect to the help variable and different ranges of possible scores: which were not observed in the other methods. The 123-method A showed a good result with respect to reliability, in contrast with the 123-method B, which showed a low value of Cronbach's alpha. At the 123-method A, the best p-value, the highest mode of the p-value, in combination with a high score of reliability were observed. The 123-method A was selected for the large sample.

Second, we determined what methods could be best applied to identify *action* of the pharmacy manager in the pharmacy mix for a large sample. The results of the method for the pharmacy manager did not correspond completely with the results of the methods for pharmacy personnel and for the researcher. In addition, the calculated clusters of the pharmacy personnel did not correspond with the 'natural clusters'. A natural cluster would be observed whenever the pharmacy personnel of one pharmacy were in the same calculated cluster. However, the respondents of the same pharmacy were never observed in one calculated cluster. This could cause problems with the interpretation of the results if this method were to be applied to a large sample. In addition, in the method for the pharmacy personnel a majority of all cases were placed in a cluster which did not correspond with the rough clusters. All cases of the method for the pharmacy manager and for the researcher corresponded to the rough clusters. The method for the pharmacy manager was the main instrument for measuring *action*. Previously, it was expected that the problem of social desirability could be at stake; the pharmacy manager selecting questions in accordance with correspondence. Additional methods were selected for the pharmacy personnel and for the researcher in order to avoid this problem. But if *non*-correspondence was observed, it was assumed that no further comparisons were required for the large sample. Within this pilot study it was illustrated that with a sole measurement of the function of pharmacy manager *non*-correspondence was observed. In addition, there were no interpretation problems with respect to the 'natural clusters'. The method for the pharmacy manager was selected for the large sample.

We agreed to determine which combination of methods could be best applied to identify the correspondence between *thought* and *action* of the pharmacy manager in the pharmacy mix for a large sample. In keeping with the outcomes mentioned above, a combination of two methods was selected. *The method for **thought** will be the 123-method A, and the method for **action** will be the method for the pharmacy manager.*

Additionally, the correspondence of 24 pharmacy organizations to their own pharmacy mix was determined. The pilot showed that the combined use of the two selected methods resulted in *non*-correspondence between *thought* and *action* of the pharmacy manager for most cases. The observed correspondence for the complete ranking was two (case 9 and 14) out of 24 pharmacies by using the 123-method A and the 123-method B for *thought* and the method for the pharmacy manager for *action*. If all results are compared it can be concluded that the 'complete' ranking (for example, product-process-customer) for *thought* and *action* corresponded in only two (product) cases; the 'double' ranking (for example, product-process) corresponded in the same two cases also; and the 'single' ranking (for example, product) corresponded in 11 cases, four of which originated from a comparison of 123-method A and the method for the pharmacy manager. In addition, results of VNA/SAL pharmacies were compared with the results of additional pharmacies. *Non*-correspondence was observed for 11 VNA/SAL pharmacies and 11 additional pharmacies. As a logical consequence, correspondence between *thought* and *action* was observed for one of the VNA/SAL pharmacies and for one of the additional pharmacies. The results of the 'single'-ranking analysis showed that the main focus for *thought* was related to actions directed at customer and product. In the method for *action* the emphasis of the actions was mainly directed at product and process.

In summary, the 123-method A was selected to represent *thought*, and the method for the pharmacy manager was selected to represent *action*. Now that the methods for the point of departure have been determined, we might as well apply these methods to a large sample. In the next part of *phase one*, **chapter 5**, we will apply these methods to a random sample of the population of Dutch community pharmacy managers. The suggested random sample size is 300 community pharmacy managers. You are reminded that in *phase two*, **chapter 6**, the organizational change to the customer mix will be studied. In **chapter 2**, we also decided to study the role of the support structure of stichting VNA and SAL Apotheken in this organizational change. Consequently, some VNA/SAL pharmacies were selected and added to the large random sample.