PERGAMON

Social Science & Medicine 54 (2002) 1225–1241

SOCIAL SCIENCE & MEDICINE

www.elsevier.com/locate/socscimed

Construction of a GP integration model

R. Batterham[a], D. Southern[b], N. Appleby[b], G. Elsworth[c], S. Fabris[c], D. Dunt[a,*], D. Young[b]

[a] *Centre for Health Program Evaluation, Department of General Practice & Public Health, The University of Melbourne, P.O. Box 477, West Heidelberg, VIC 3081, Australia*

[b] *Section of General Practice, Department of General Practice & Public Health, The University of Melbourne, Australia*

[c] *Centre for Programme Evaluation, Department of Education Policy & Management, The University of Melbourne, Australia*

Abstract

There are frequent calls to improve integration of health services, within and between primary and secondary care sectors. In Australia, general medical practitioners (GPs) are central to these endeavours. This paper aims to better conceptualise GP integration and to develop a model and index based on this.

A conceptualisation of integration is proposed based on integration fundamentally as an activity or process not structure. Integration process is the frequency and quality of episodes of information exchange involving the GP and another practitioner or patient and aimed at fulfilling the objectives of the health care system with regard to patient care. These are both direct responses to structural forces and emergent GP capacities and dispositions.

The content of this typology was studied using Concept Mapping in 11 groups of GPs, consumers and other practitioners. Clusters of related statements within thematic domains were used as the basis for a provisional model. This was tested using confirmatory factor analysis in a data set derived from a national probability sample of 501 GPs.

Some re-specification of the model was necessary, with three integration process factors needing to be subdivided. One factor congeneric model assumptions were used to identify the constituent items for these factors. The result was a model in which 50 items measured nine integration process factors and 20 items measured five enabling factors. Two distinct but correlated higher order factors, relating to individual patient care and public (or community) health—in contrast to a single higher order factor for integration—were identified. The re-specified model was tested with a new sample of 151 GPs and exhibited strong psychometric properties. Reliability and validity were acceptable to this stage of the indices' development. Further testing of the index is necessary to demonstrate factor invariance of the indices in other contexts as well as their utility in cross-structural analysis. That said, the indices have immediate uses.

Keywords: GP integration; Concept Mapping; Index construction

Introduction

Complex, multi-dimensional social concepts or phenomena deemed to have positive impacts on community health (such as 'community empowerment', 'service responsiveness') are in widespread use. Surprisingly, there have been few attempts to examine their meaning closely or to identify their dimensions. Perhaps their very complexity deters such efforts. However, study of the concept would have value in demonstrating that it has, or does not have a stable reality, judged by agreement about its meaning and existence of stable patterns of observations implying such a construct.

The complex, multi-dimensional concept studied in this paper is 'integration of general practitioners with the health care system'. In the words of the emergent realists, the paper investigates whether 'GP integration' is something 'real' not merely a convenient shorthand

*Corresponding author. Tel.: +61-3-9496-4433; fax: +61-3-9496-4424.

E-mail addresses: d.dunt@gpph.unimelb.edu.au (D. Dunt).

PII: S0277-9536(01)00092-2

term to describe a class of programs (Cronbach, 1989). Given the lack of preexisting theory, its methodology is exploratory and iterative and inevitably involves judgements of the investigators which are influenced by their presuppositions and values. There is however no alternative approach save abandoning the task.

In this paper, Concept Mapping is used to specify a theoretical model, which can be subsequently tested by confirmatory factor analysis, using structural equation modelling (SEM) in an appropriate data set. This hypothesis generation and testing approach contrasts with the traditional hypothesis generating step only of exploratory factor analytic (EFA) approach (McArdle, 1996). In the latter, a convenience sample of subjects or researchers is typically used to generate statements or questions for which no model is specified. Factors are subsequently derived from these statements in a data set using EFA.

Concept Mapping is a nominal group technique that permits the rapid development of a conceptual framework for any given topic (Trochim, 1989a). It is expressed entirely in the language of the participants and yields a graphic or pictorial product which simultaneously shows all major ideas and their interrelationship. To construct the map, ideas (statements) have first to be described or generated and the interrelationship between them articulated. Multi-variate statistical techniques are then applied to this information and the results are depicted in map form.

The paper describes a measurement model of GP integration as distinct from a structural causal model which would investigate the relationship between the whole (or indeed parts) of the GP integration model with other concepts and variables. It is necessary for the measurement model to be tested and refined before detailed exploration of causal models is possible.

The significance of GP integration within the health care system

The role of the general medical practitioner (GP) in the Australian health care system

The Australian health care system is a mixed public/private system. It is further fragmented by a federal system of government where the Commonwealth Government's role is to fund public health services within a large number of programs and to underwrite private care, and the States and Territories' role is to be responsible for the delivery of public health services. Universal medical insurance covers medical practitioner services but fee rebates for these services are only obtainable if the patient is referred by a GP. GPs usually operate in independent private businesses and are paid on a fee-for-service basis either by the patient or through reimbursement directly from the government insurer. As a result of their gatekeeper role GPs are central to the health system. The establishment of local area organisations called Divisions of General Practice has provided some support to GPs in these and other newer roles (General Practice Changing the Future Through Partnerships, 1998).

The necessity for integration

Integration of primary health care with the remainder of the health care system has been identified as an important reform imperative in quite different health care systems. A variety of different ways of achieving this have been proposed. At the point of contact with the patient, case management approaches have become widespread (Falik et al., 1993). Other approaches have focussed more on systemic reforms such as through the Health Maintenance Organisation in the US and GP fundholding in the UK. In systems with GPs, their integration with the remainder of the system has been central to these reforms. This is because GPs are typically both the first and continuing contact point for patients seeking health care and the gatekeepers of the health care system.

They can become involved in service integration through both point of contact and systemic reform approaches. For example, a major service reform aimed at service integration has recently been instituted (Common Wealth Dept. of Health and Community Services, 1995). This is a National Trial of Coordinated Care, based around the GP as Care Coordinator (sometimes in conjunction with a generic case manager). The Trial involves the development of individualised care plans by GPs for patients with complex care needs and which aim to harness a different mix of service providers in fulfillment of these needs.

In the UK, a series of recent white papers presented opportunities for general practice to extend and develop their services to patients. It has been argued that these could enhance professional roles within practices and lead to new partnerships with secondary care, community health services and social services (Kendrick & Hilton, 1997). These opportunities may be reinforced under the 1997 Primary Care Act and fundholding arrangements involving the new Primary Care Groups (Dixon, Holland, & May, 1998). Other common GP integration initiatives include obstetrics, diabetes and mental health Shared Care, GP hospital liaison schemes and Hospital in the Home arrangements (Harris, Fisher, & Knowlden, 1993; Montalto & Dunt, 1993). IT strategies linking GPs to other providers and hospitals are becoming increasingly common (Liaw, Lawrence, & Rendell, 1996).

Whether these general practice integration initiatives really promote better use of resources, and lead to better, long-term patient health outcomes remain largely unanswered. Even assuming that they do, it is unclear what form these integrative strategies should take. A constraint on their evaluation is that no valid or reliable index is available for measuring the level of GP integration achieved in the particular setting being evaluated. This may mean that in studies of the concept and its effects, the organisational arrangements believed to produce or enhance it, rather than the concept itself are studied. A study of GP integration that employed a randomised control trial to examine the impact of obstetric shared care on health service usage costs would be an example of this. A GP integration index could be used to interpret program effects—why a particular arrangement did or did not produce expected impacts or health outcomes. An understanding of this may suggest how the program can be improved in the future, and with it, the outcomes of that program. More generally, a study of the concept of GP integration through the attempted development of a relevant index makes a contribution by demonstrating whether or not the concept really exists and has utility in understanding the important process of health care systems.

Developing a conceptual schema to describe GP integration

To begin, the Macquarie Dictionary (1997) defines integration as 'the act of bringing together parts (into a whole)'. Coordination, by comparison, is defined as 'the act of arranging according to a plan especially one drawing together a number of different events, organisations, people, etc.' and 'working together smoothly in combination'.

Descriptions of integration of health services (not GP integration) that exist in the research literature are primarily used for purposes of classification. For example the concepts of 'vertical' and 'horizontal' integration merely describe the sorts of agencies that are in relationship (Dixon et al., 1998). Øvretveit (1993) proposes a multi-dimensional typology of coordination (if not integration) of care that encompasses both structure and mechanism. Three structures with the potential to create coordination—bureaucracy, markets and associations—are described. Within each, there exist a number of possible models for cooperation between purchasers, between purchasers and providers, and between providers. A number of different sorts of care teams are described that differ with regard to their degree of permanence and formality, their leadership, and their degree of freedom to make decisions. A functional taxonomy of the purposes for which teams meet (including reception, assessment, allocation, intervention, and review) is also presented. No description or definition of GP integration beyond the classificatory systems, such as those above, has been encountered.

A wider study of the nature of social organisations was more fruitful. According to the Cultural—Historical Activity System model of Engestrom (1998), human activity is a collective, systemic formation that has a complex mediational structure. An activity system involves objects and produces actions and is realised by means of actions. However, activity is not reducible to actions. Actions are relatively short-lived and have a temporally clear-cut beginning and end. Activity systems evolve over lengthy periods of socio-historical time, often taking the form of institutions and organisations. Collective activity is connected to object and motive, of which the individual subjects are often not consciously aware. Individual action is connected to a more or less conscious goal.

The Activity System model involves subject, object, outcomes, instruments, community, division of labour and rules and can be illustrated in terms of a primary care clinic, as follows The *subject* (physician) is chosen as the point of reference. The *object* (patients) refers to the 'raw material' or 'problem space' at which the activity is directed and which is moulded and transformed into *outcomes* (intended improvements in health plus other unintended outcomes) with the help of physical and symbolic, external and internal mediating *instruments* (X-rays, medical records plus partially internalised diagnostic and treatment-related concepts and methods). The *community* (staff of the clinic) comprises multiple individuals and/or sub-groups who share the same general object and who construct themselves as distinct from other communities. The *division of labour* determines the tasks and decision-making powers of the physician, the nurse, the nurse's aide, and other employee categories. Finally, the *rules* regulate the use of time, the measurement of outcomes, and the criteria for rewards. An activity system also interacts with a network of other activity systems, e.g., it receives rules and instruments from certain activity systems such as management.

An activity system is always heterogeneous and multi-voiced. Different subjects, due to their different histories and positions in the division of labour, construct the object and the other components of the activity in different, partially overlapping and partially conflicting ways. There is constant construction and re-negotiation within the activity system, e.g., when practitioners of a medical clinic, using experiences from other clinics, design and adopt a new model for their work that corresponds to the ideals of a more holistic and integrated medicine. The new ideas may be formally implemented, but they are internally resisted by the vestiges of the old activity. Tasks are reassigned and redivided, rules are bent and reinterpreted. Beyond this, contradictions may sometimes occur between different elements of the system or between the system and other systems. This introduces disequilibrium and change into the activity system. In all these different circumstances, coordination between different subjects and their versions of the object must be achieved to ensure continuous operation.

Applying these insights to the concept of GP integration, it is clear that process is the primary dimension and structure only secondary. The focus is on what GPs do and how they interact with other providers and consumers in the health care system. *Organisational structures* are developed to improve integration process. Focussing on integration as *process*, it is useful to consider what the unit, or single episode, to be termed the *integration event,* might be. Based on the above, it may be provisionally defined as:

> an episode of information exchange involving the GP and another practitioner or patient that establishes a mutuality of understanding or intent aimed at fulfilling the objectives of the health care system with regard to patient care.

Considering integration as *structure*, organisational *resources* or *incentives* can be brought to bear to influence integration events in a straightforward fashion. Other organisational arrangements such as *regulations* or *sanctions* can do this also though sometimes in a more indirect way.

This leads to a definition of GP integration as process:

> the pattern of execution (frequency, nature and diversity) of integration events as represented in the characteristic practices of particular GPs.

It also leads to a definition of integration as structure:

> the organisational inputs (resources, incentives, regulations and sanctions) that promote and support the integration process (as defined above).

It is proposed further that if these structural arrangements continue and are favourably received, they are likely to have impacts on the integration process beyond the immediate stimulus on the GP to engage in a particular integration event. These involve the GP's 'capacity' to integrate in the sense used by Bandura (1982) in his concept of self-efficacy. They should also impact on their 'predisposition' to integrate, e.g., to engage more frequently in integration events in the future. An enhanced capacity and predisposition to integrate could conceivably lead to the GP engaging in new forms of integrative activity.

In addition, GP integration involves not only the GP but also the patients and other practitioners who have contact with the GP. The integration structural forces noted above will impact not only on GPs but also their patients and these other practitioners both directly as well as on their capacity and predisposition to integrate. These elements of this model are illustrated in Fig. 1 below.

To what extent the capacities and predispositions of these practitioners, say renal physicians or occupational therapists, differentially impact on the GP bringing forth a variable level of GP integration behaviours in different contexts is at this point an open question to be empirically determined.

Thus this initial typology is contentless but provides a starting point for proceeding with the next stage of index construction—Concept Mapping. It is clear though that there will be a focus on the cognitive and affective characteristics (capacity, predisposition) of individuals rather than formal organisational structures such as might exist in hospitals. In addition, only GPs (with the possible exception of a few other practitioners who work very closely with GPs) would be able to report accurately on these GP characteristics and therefore on

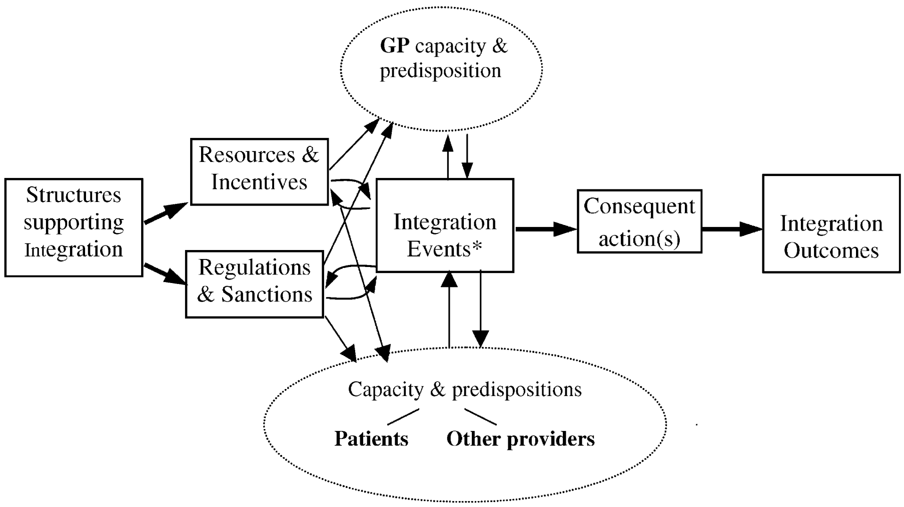


Fig. 1. Integration pathway.

the concept of GP integration itself. For these reasons, enquiries in the Concept Mapping groups described below began with the broad question to GPs and stakeholders:

> *Thinking as broadly as possible, generate statements that describe the role and circumstances of a GP who is well integrated into the health care system.*

Specifying a model of GP integration through application of Concept Mapping techniques in small groups

Concept Mapping was used to specify the model of GP integration to be subsequently tested using SEM. It is a nominal group technique that permits the rapid development of a conceptual framework for any given topic (Trochim, 1989a). It is expressed entirely in the language of the participants and yields a graphic or pictorial product which simultaneously shows all major ideas and their interrelationship. To construct the map, ideas (statements) have first to be described or generated and the interrelationship between them articulated. Multi-variate statistical techniques are then applied to this information and the results are depicted in map form.

The use of Concept Mapping for the identification of clusters (or groups of related statements) and specifying the nature of their interrelationship, within a nominated topic area, is well established (Galvin, 1989; Markham, Mintzes, & Jones, 1994; Marquart, 1989). Concept Mapping can be used to better understand complex or multi-faceted goals and objectives of programs by defining their component parts. This can be done in a variety of ways in the context of both planning and evaluation, one of which is to develop question inventories for either one or all of these component parts. Trochim (1989b) emphasises that Concept Mapping should be thought of as a process of theory development, which then needs to be tested using various empirical techniques.

Concept Mapping methods

The methodology outlined here has been described more fully elsewhere (Southern, Batterham, Appleby, & Young, 1999).

A purposeful sampling technique was used to obtain 11 suitable groups of like individuals including four GP groups with different characteristics and seven groups of non-GP health providers, managers and consumers. These were drawn from three Australian States (Victoria, Western Australia and Queensland). Groups were asked to generate statements describing the role and circumstances of a GP who was well integrated into the health care system. Participants sorted items so generated on the basis of their perceived inter-relatedness. Participants rated each statement in terms of its 'importance to patient care' and 'present level of attainment'.

Statements were represented in the form of Concept Maps using non-metric, multi-dimensional scaling based on the sorted statements of each group member and hierarchical cluster analysis. Concept Mapping software was used to perform this multi-dimensional scaling. A two-dimensional map of points representing a binary similarity matrix of the generated items was created. The points were partitioned using hierarchical Concept Mapping into groups or clusters and overarching domains of clusters using Ward's algorithm (Everett, 1980). The location of clusters vis a vis other clusters on the maps depends on the similarity of the statements on which they are based. Clusters with conceptually similar statements are located more closely together than ones with less similar statements. Participants were presented with these maps. They then labelled both clusters and thematic domains formed by related clusters. An example of a cluster map is shown in Fig. 2.

An ordered meta-matrix (Miles & Huberman, 1988) that displayed clusters and thematic domains from the concept maps of the 11 groups was then constructed so as to identify the provisional factors to be tested using SEM. This made apparent the extent of variation in naming practices between the 11 groups. It was nevertheless possible to identify these provisional factors by inspection of the ordered meta-matrix in conjunction with a spatial interpretation of the 11 Concept Maps. The latter ensured that provisional factors so identifed were distinct or non-overlapping, i.e., distant from each other on the maps. When a decision was required whether to include a factor or not, an inclusive approach was used so as to not exclude at this stage potentially important information about GP integration.

Each statement generated in the Concept Mapping step sessions was independently assigned to one of these provisional factors by two members of the study team (DS and NA). Differences in assignment were discussed and resolved. A review of these statements by DS, NA and a GP reference group indicated that a large number of these were unsuitable for survey purposes because they were duplicates of other statements, contained more than one idea, were ambiguous or difficult to interpret.

For a variety of reasons, the model specified by this Concept Mapping approach is likely to be approximate and to requires some re-specification at the SEM stage. For example, clusters generated are very sensitive to the clustering algorithm in the hierarchical cluster analysis program. In addition, at the consolidation stage, subjective decision making by the research group was unavoidable. As noted by Shumaker and Lomax (1996),

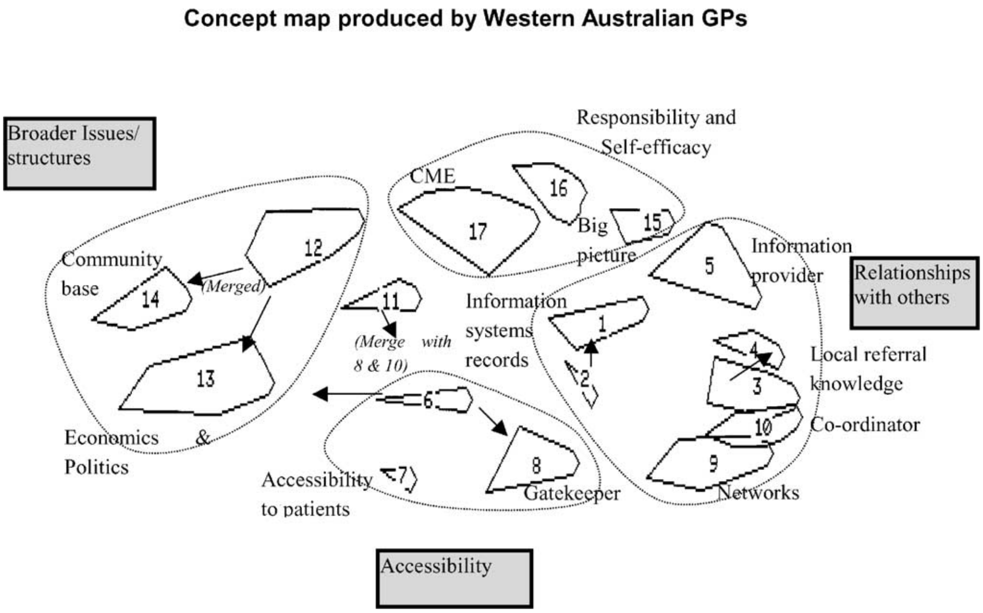


Fig. 2. Concept Map produced by Western Australian GPs.

this procedure, called a 'specification search', is more common than in the application of SEM. Its purpose is to refine the model at the margins rather than develop a whole new model.

Concept Mapping results

Full results with regard to Concept Maps, consolidation matrix and final list of constituent statements for each factor have been described in full elsewhere (Young et al., 1999). Fig. 2 shows one of the Concept Maps, in this case that produced by a group of GPs in Western Australia. The numbered clusters each contain a number of statements, the ratings for each cluster are the average ratings of the statements within it. In addition to labelling *clusters* based on the ideas common to the statements, participants were asked to look for any broader groupings or *domains*, ideas that might be common to a number of clusters. These are marked on the map with dotted lines and labelled in the shaded boxes. Generally speaking clusters on the right relate to local relationships while those on the left relate to the broader contextual determinants of integration. The top relates to GP characteristics while the bottom mostly relates to patient experiences. This partitioning of the space in terms of stakeholder groups (with whom GPs are integrating) was common to many of the maps, a finding that was important in hypothesising a set of factors for a model to measure integration.

Table 1 lists the clusters for four of the 11 groups in descending order of importance.

While some differences emerged, GP and non-GP groups broadly identified similar process and structural dimensions of GP integration. Clusters relating to the services provided to patients and GP characteristics rated highest on importance. Those relating to the broad policy context in which GPs work were considered to be most problematic both in terms of whether they were happening and the difference between this and their importance.

Inspection of the ordered meta-matrix in conjunction with a spatial interpretation of the 11 Concept Maps indicated that GP integration could be expressed in terms of five important stakeholder groups and that, in each case, there were identifiable major functions and essential characteristics of integration.

The stakeholder group and the identifiable characteristics of integration for that group are as follows:

1. *Patient group (Holistic and flexible)*. Integration with patients was conceptualised as creating a partnership in the process of care in which the GP acted as a guide to the care system but was all the time sensitive to the holistic (physical, emotional, social and spiritual) needs of patients. To facilitate this partnership it was suggested that GPs need to be flexible in how they work with patients.
2. *Community (Community health)*. It was considered that the well integrated GP would have a substantial role in identifying health problems in a community

Table 1
Clusters identified by four of 11 groups listed in order of importance

Community health centre GPs	*Queensland rural GPs*
Patient knowledge 1	Communication skills
GP as educator	Infrastructure and resources
General community links 1	Broader professional interests/appointments
Core activities	Proactive and evolving approach to skill development/application
Patient knowledge 2	Resource utilisation and awareness
General community links 2	GP as educator
Education of GP	Diverse patient management and treatment paradigms
Community roles	Holistic patient treatment
Constraints on GP role flexibility	Patient health care coordinator
Systemic issues	GP's personal life-style management
Aged care assessment teams	*Consumers*
Sensitivity and respect for all players	Patient information and consent
Appropriate referral	Role identification
Collaboration	Respects patient rights
Personal and professional attributes	Sensitive to whole patient
Awareness and understanding of service providers	Translators
Responsiveness to other authorities	Sensible about role and limitations
Case management awareness	Role in continuity of care
Health system awareness	Multi-disciplinary team
Policy planning role	Professional development and support
Patient education/Health promotion	Aware of community and educates community
	GP's own well-being

and acting to meet those problems. Their ability to fulfil this role is in part dependent on the connections they maintain in the community.

3. *Local service providers (Teamwork)*. Integration with other service providers was considered to have a number of levels from informal liaison and simple information exchange right through to collocated interdisciplinary teams and shared care arrangements.
4. *Hospitals (including specialists)/(Care-coordination)*. It was considered essential that GPs have a role throughout the whole hospitalisation experience of their patients from preadmission assessment and preparation of the patient, during the inpatient period and after disharge.
5. *The health system (Health planning)*. It was considered that well integrated GPs have a role in both implementing and influencing health policy. This level of integration in part depended on the extent to which GPs were effectively joined together to have a common voice.

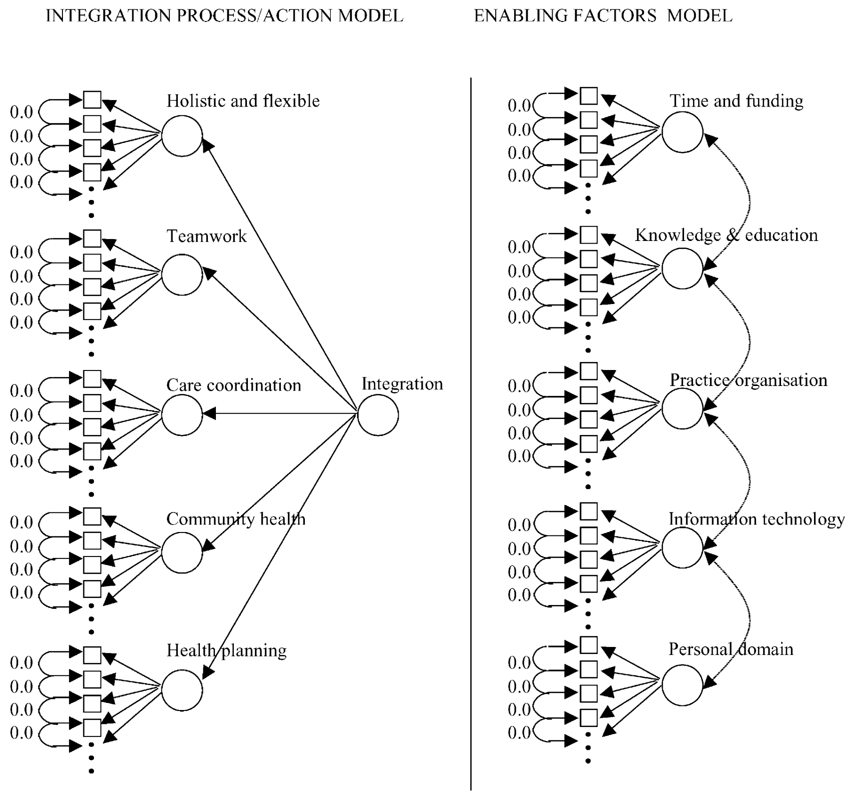
A large number of diverse statements relate to three of these factors (integration with patients, other local providers, and hospitals and specialists). This suggested that further sub-factors could be embedded within each of these three factors.

Five influences on these characteristics of GP integration could also be identified. These were:

1. *Time and funding*: particularly referred to the need to have flexible funding arrangements to support a range of integration activities.
2. *Knowledge and education*: referred to the need for GPs to be involved in continuing medical education, to keep their knowledge about local services up to date and to have resource materials readily available.
3. *Personal attributes*: characteristics of the GPs important for integration.
4. *Practice organisation*: covered issues such as the record system used in the practice, resources available and, initially, information technology.
5. *Information technology*: this was initially viewed as part of 'practice organisation' but at the suggestion of a GP reference group, it was split off to form a hypothesised factor on its own because of its perceived importance and independence from other aspects of practice organisation.

These 10 provisional factors then formed the basis for the model tested using SEM and described in the next section—see Fig. 3.

Seven hundred and fourteen statements were generated in the Concept Mapping step sessions and were assigned to one of these 10 factors as described above.



The conventions used in the diagram are standard except that the dots below the items for each factor suggest that the final number of indicators is still to be determined. Zero correlations between the unique variances of the items show that the assumptions of the congeneric model operate. The curved arrows between the enabling factors indicate that they are free to covary although only the arrows between adjacent factors are drawn in. All other undrawn relationships (eg between the unique variances of higher order factors, factors, and other items) are postulated to equal zero.

Fig. 3. Model of GP integration derived from Concept Mapping.

After reviewing the suitability of these statements for survey purposes, 114 statements remained and were subsequently used in the GP mail questionnaire. The questionnaire is available from the research group on request.

Confirming and re-specifying GP integration models using SEM in a national GP survey data set

Two variants of the hypothesised models for the integration process factors and the enabling factors were proposed. The simpler without higher order factors is shown in Fig. 3 above. A more complex variant of the integration process factor model (not diagrammed) encompassed the simple variant plus three mediating higher order factors between the higher order factor *integration* and the five individual factors. These three higher order factors represented several though not all of the five individual factors. These mediating factors were nominated as being a *GP patient care* factor, a *collaboration with other health care professionals in patient care* factor and a *broader role with other organisations and community* factor.

The confirmation and re-specification of this model using SEM proceeded through the following steps. Two national GP sample surveys to create calibration and validation data sets were conducted. Individual factors and items within integration process and enabling factor models were tested and re-specified where necessary using data from the first survey (calibration sample). Testing (with re-specification as necessary) of a complete measurement model for both integration process and enabling factor models, first in the absence and then in the presence of higher order factors, was also conducted on the calibration sample. Finally, confirmation of all individual factor and multi-factor measurement models was carried out on data from the second (validation) survey.

Conduct of national GP sample surveys to create calibration and validation data sets

Two separate national surveys of GPs were undertaken. Only GPs substantially involved in general practice were deemed eligible (this was judged on the basis that they had provided at least 1500 billable services in the most recent 12 months for which data were available).

The *first (calibration) sample* was surveyed to test and re-specify the models as necessary. GPs were drawn from two sources—900 GPs randomly selected by the Australian Health Insurance Commission (HIC), stratified by gender and metropolitan/rural/remote status, and one representative GP chosen by each of the 118 Divisions of General Practice, each representing regional groupings of GPs throughout Australia. GPs were asked to report on their actual behaviour in their clinics. Questions consisted of the 114 topics/items generated by the Concept Mapping exercise. Five-point Likert scales were used despite a preference for six-point, forced choice scales since GPs in a pilot survey strongly felt that the forced choice format could reduce response rates. The questionnaire also contained questions on the actual integration events undertaken in the past month and on the GP's socio-demographic and practice characteristics.

The metropolitan, male GPs in the HIC sub-sample were over-sampled 3 : 1 as their response rates are usually lower than other groups. The response rate was 251 (49.5%), being similar for most sub-groups but substantially higher for female, rural GPs (64%).

The *second (validation) sample* of 151 GPs was surveyed to confirm the re-specified models. The 'target' ' sample comprised three representative GPs chosen by each of the 118 Divisions of General Practice plus 20 other GPs known to the research group. These three representative GPs were to include one male and one female GP as well as one GP closely involved and one not so closely involved in Divisional activities. The response rate was 39.7%. Questions in the validation sample questionnaire consisted of only the 70 of the 114 items that the analysis of the calibration sample data set had indicated as useful. Six-point rather than five-point Likert scales were also used.

Testing and re-specification of individual factors and selection of items

Factors

Method. SEM analyses were carried out using PRELIS (Jöreskog & Sörbom, 1988) for structuring of the data, and LISREL 7.0 (Jöreskog & Sörbom, 1989) for analysis of single. The fit of each factor to constituent items was assessed using the assumption of the congeneric model, i.e., that all covariance must be due to the one underlying factor. Analyses were based on polychoric correlations, thereby taking into account the ordinal nature of the response scales. Where missing values for data existed, group means were inserted.

Identification of the optimal set of items for each factor in both models was assessed using seven statistics reported by LISREL 7.0. Four are based on comparing the covariance matrix predicted by the model with the covariance matrix represented by the actual data (χ^2, goodness-of-fit index (GFI), adjusted goodness-of-fit index (AGFI) and root-mean-square-residual, (RMSR)). Modification indices (MI) suggested where the single factor assumption might break down by indicating the amount by which the χ^2 would improve if two items were allowed to covary uniquely. The coefficient of determination (Rc) measured the proportion of factor variance accounted for by a weighted linear combination of the item scores. It represents the success of the scale in measuring the factor and provided one estimate of reliability.

In assessing and re-specifying the models, GFIs and AGFIs above 0.9 and RMSRs less than 0.1 were considered desirable. MIs greater than two were examined, though with only much larger values leading to re-specifications. Rc values above 0.9 were considered desirable though values above 0.8 were accepted.

Factors with sub-optimal properties were further examined to identify if they contained sub-factors with improved properties. Traditional exploratory factor analytic techniques were used to assist in this process.

Results. Only one of the five proposed integration process factors had psychometrically strong properties and could be confirmed. Four of the integration process factors therefore required re-specification. Significant problems involving poor AGFIs were identified for Factor 1 (Holistic and flexible), Factor 2 (Teamwork), Factor 3 (Care coordination) and Factor 5 (Health planning). Three sub-factors (Holistic patient care, GP Flexibility, Provision of information to patients) were identified for Factor 1 (Holistic and flexible), two (Attitudes to teamwork,Liaison) for Factor 2 (Teamwork) and two (Care coordination (non-hospital), Hospital role) for Factor 3 (Care coordination).

The psychometric properties of these seven sub-factors were tested and were much improved on the original three factors of which they formed part. Accordingly, the original integration process factor model was re-specified with three of the original factors being replaced by seven. In addition, no clear sub-factor structure could be identified for Factor 5 (Health planning) and the problems with this factor were resolved by narrowing the concept and reducing the number of items.

Four of the five enabling factors had psychometrically strong properties and could be confirmed. Factor 7 (Time and funding) had problems associated with its poor reliability. These were considered to be not serious enough to necessitate modifying the factor at this stage, though this did in fact occur at the measurement model stage. Subject to this caveat, the enabling factors were confirmed up to this stage.

Items

Method. For each factor a set of items with optimal psychometric properties was identified through a process of deleting items iteratively. In doing so, it should be noted that retention of items usually improves the reliability of the factor (measured here by Rc) but may lead to the introduction of sub-factors decreasing the fit of a single factor congeneric model (best measured by the AGFI and the RMR). Given these mixed effects, the item set chosen was not necessarily associated with optimal values for all statistics, but rather represented optimal values across the different statistics, balancing these different effects. Items which poorly represented the factor as indicated by a low squared multiple correlation (SMC) were generally deleted. While SMCs of more than 0.6 were preferred, in one case an item with a SMC a little over 0.4 was retained because it substantially improved the properties of the scale as a whole.

Results. Items were deleted from all but two of the 14 integration and enabling factors. As a result, the final version of the questionnaire, which was sent to the validation sample, consisted of 70 items (50 associated with the integration process model and 20 with the enabling factors model) reduced from 114. This number of questions produced a questionnaire of suitable length, which is available from the authors on request. This set of items was subjected to one further review at the measurement model stage (see below) at which point two further of the 20 enabling items were excluded.

Confirmation of a measurement model for both integration and enabling factor models

In the absence of higher order factors

Method. Separate confirmatory factor analyses of the whole integration process model and the enabling factor model based on these revised factors and items were conducted using EQS 5.6b (Bentler & Wu, 1995). This was done to identify any cross-loading items and to determine whether the model and factors required any further re-specification. Only when this was established was a model including higher order factors tested.

Twelve Fit statistics are reported by EQS 5.7b (Bentler, 1995). The Bentler–Bonnet non-normed fit index (NNFI), the comparative fit index (CFI), the root mean square error of approximation (RMSEA) and the upper confidence level of the RMSEA are considered to be most important of these (Bentler, 1995). These indices will be reported through the remainder of the paper. As discussed by Pedhazur and Pedhazur Schmelkin (1991), the criteria for an acceptable fit depend on many circumstances, although various authors suggest RMSEA of less than 0.05 or its upper 90% confidence limit of less than 0.08 as useful rules of thumb. A NNFI or CFI of less than 0.90 could be taken to indicate a poor fit. Many authors have noted that χ^2 is invariably statistically significant in large models with large samples. Under these circumstances, the RMSEA is considered a better measure of the extent to which the reproduced covariance matrix differs from the sample covariance matrix.

The analysis was conducted on the covariance matrices for both the calibration and validation samples. Given the multi-variate kurtosis present in the data set, elliptically reweighted least squares (ERLS) estimation was used in preference to maximum likelihood (ML) estimation which makes strong assumptions about multi-variate normality (Sharma, Durvasula, & Dillon, 1989).

Initially the model was constrained such that there were no inter-item correlated errors and no items loaded on more than one factor. In accordance with the principles of confirmatory factor analysis, these constraints were relaxed only if the correlated errors or cross-loadings were theoretically plausible.

Results—Integration process factor model. The model was fitted after the following re-specifications. Four cross loadings of items on more than one factor were identified and accepted with the relevant items being allocated to the scale where they had the strongest loading. This resulted in two items being reassigned (Item 10 in Factor 9 Health planning and Item 12 in Factor 4 liaison were reassigned to Factor 8 Community health). An inter-item correlated error between two items, one in the Health planning factor, the other in the Community health factor, was also allowed (see Table 3). These changes to the model were considered to be minor and theoretically supportable. The fit of the re-specified model to the data was very good (see Table 2), judged by the criteria outlined above.

Results—Enabling factor model. This model could only be fitted after the following re-specifications—see Table 2. Factor 2 (Time and funding), which had already demonstrated poor reliability as a single factor, needed to be revised with two of the four items being removed. As the fit of the resulting two-item scale cannot be tested, it cannot be regarded as a valid factor. As an 'indictor' of 'Time and funding' it is superior to the two separate items and is retained in the model in this form at this stage. Second, constraints on two item crossloadings were also relaxed—Items 41 and 44 each crossloaded on two different factors.

Table 2
Fit statistics for the integration process and enabling factor measurement models

Fit statistic	Calibration sample	Validation sample	Two-group analysis
(a) *Model for re-specified integration process factors*			
χ^2, df, p	1604, 1134, <0.001	1590, 1134, <0.001	3394, 2359, <0.001
χ^2/df	1.4	1.4	1.4
NNFI	0.97	0.93	0.95
CFI	0.97	0.94	0.95
RMSEA (±90% CIs)	0.041 (0.036–0.045)	0.049 (0.043–0.054)	0.033 (0.030–0.035)
(b) *Model for re-specified enabling factors*			
χ^2, df, p	261, 122, <0.001	178, 124, <0.001	468, 259, <0.001
χ^2/df	2.1	1.4	1.8
NNFI	0.95	0.89	0.94
CFI	0.96	0.91	0.95
RMSEA (±90% CIs)	0.07 (0.06–0.08)	0.05 (0.03–0.07)	0.04 (0.04–0.05)
(c) *Model for re-specified integration process factors (incorporating higher order factors)*			
χ^2, df, p	1701, 1159, <0.001	1681, 1159, <0.001	3466, 2375, <0.001
χ^2/df	1.47	1.45	1.46
NNFI	0.96	0.92	0.95
CFI	0.96	0.93	0.95
RMSEA (±90% CIs)	0.043 (0.039–0.047)	0.052 (0.046–0.057)	0.033 (0.031–0.035)

Table 3
Standardised solutions for integration process factor model[a]

Factor	Calibration sample	Validation sample
Holistic patient care =	0.751 * F1 + 0.660 * D1	0.895 * F1 + 0.445 * D1
Provision of information to patients =	0.604 * F1 + 0.797 * D2	0.642 * F1 + 0.766 * D2
GP flexibility =	0.955 * F1 + 0.297 * D3	0.933 * F1 + 0.359 * D3
Liaison =	0.408 * F1 + 0.523 * F2 + 0.593 * D4	0.379 * F1 + 0.491 * F2 + 0.663 * D4
Attitudes to team work =	0.948 * F1 + 0.319 * D5	0.836 * F1 + 0.549 * D5
Hospital role =	0.735 * F2 + 0.678 * D6	0.284 * F2 + 0.959 * D6
Care coordination (non-hospital) =	0.958 * F1 + 0.288 * D7	0.917 * F1 + 0.399 * D7
Community health =	0.830 * F2 + 0.558 * D8	0.995 * F2 + 0.102 * D9
Health planning =	0.850 * F2 + 0.527 * D9	0.737 * F2 + 0.675 * D8

[a] F1and F2 are *Patient-care integration* and *Public health integration* higher order factors, respectively. D1, D2, etc., are designated 'disturbance' terms representing the variance in the factor scores not explained by the model. Square of the higher order factor loading represents the proportion of each factor explained by the higher order factor. All paths noted in each of the standardised models are statistically significant.

In the presence of higher order factors

It was not possible to adequately fit the model with a single higher order factor *Integration*, specified in Fig. 3. This was true whether the three intervening mediating higher order factors (described above) *GP patient care* factor, *Collaboration with other health care professionals in patient care* factor and a *Broader role with other organisations and community* factor were incorporated or not. It was possible though to successfully fit a model with two higher order factors (see Table 3), judged by the criteria set out in (i) directly above. The first *Patient-care integration* combined the first two of the hypothesised intervening factors. The second *Public health integration* was the third of the intervening factors *Broader role with other organisations and community* factor, renamed. The standardised ERLS solution for the integration process factor model confirmed the existence of these two higher order factors (see Table 3).

The two higher order factors were correlated ($r = 0.54$) and this was significant, suggesting the possible existence of a third-order factor. Such a factor was not incorporated in the model, however, since the level of correlation was not thought to be strong enough.

Table 4
Selected psychometric statistics for final integration process and enabling factors in calibration sample[a]

No.	Factor	GFI	AGFI	RMSR	FB index (calibration sample)	FB index (validation sample)
Integration process factors						
1	Holistic patient care	1.00	0.99	0.03	0.91	0.89
2	GP flexibility	1.00	0.99	0.03	0.91	0.86
3	Patient information	0.99	0.99	0.04	0.75	0.73
4	Liaison	1.00	0.99	0.03	0.71	0.67
5	Attitude to teamwork	1.00	0.99	0.03	0.85	0.82
6	Care coordination (non-hosp)	1.00	0.99	0.03	0.86	0.81
7	Hospital role	1.00	0.99	0.04	0.85	0.86
8	Community health	0.98	0.97	0.05	0.91	0.87
9	Health planning	1.00	1.00	0.01	0.89	0.80
Enabling factors						
10	Knowledge of local resources	1.00	0.995	0.02	0.88	0.83
11	Practice organisation	1.00	0.999	0.01	0.98	0.81
12	Personal domain	0.98	0.911	0.07	0.85	0.81
13	Information technology	0.97	0.855	0.09	0.60	0.69
14	Time and funding	NA	NA	NA	NA	NA

[a] The FB Index refers to the index of reliability for scales that are congeneric but not fully Tau equivalent proposed by Fleishman and Benson (1987). Fit statistics and reliability estimates are not provided for 'Time and Funding' as the two items with adequate properties are not being treated as a scale.

Confirmation of factors and measurement models in validation sample

Method. As stated, it was not possible to confirm the integration process and enabling models generated by Concept Mapping groups in the calibration sample without some degree of re-specification. Consequently, it was necessary to confirm the (re-specified) model in a new sample (the validation sample). This was done at both individual factor and overall measurement model levels. The structures (though not the factor loadings) of the re-specified models were applied to the validation sample data set. The factor loadings and the fit statistics generated were compared with those obtained from the calibration sample. This was done both by direct inspection and through the use of a two-group analysis of the unpooled data from the two samples. In this analysis all factor loadings were constrained to be equal, a constraint which would be consistent with a constant, generative mechanism underlying the factor structure in the two samples. This analysis was done for the models both with and without the higher order factors.

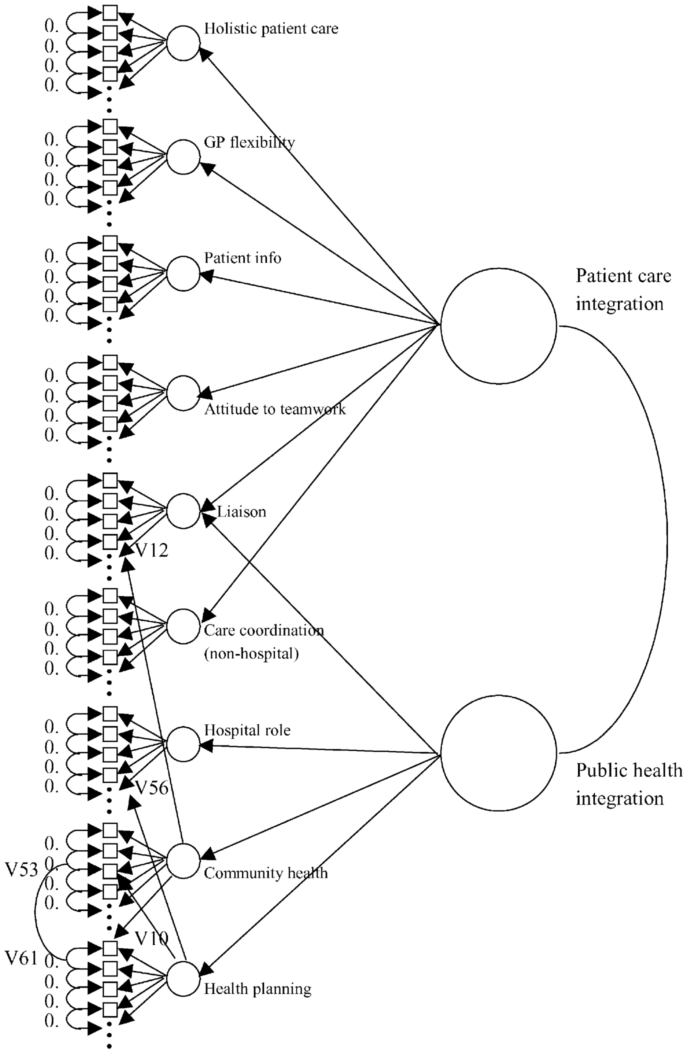
Results. Fit statistics for each sample are reported in Table 2. The model structure was confirmed for both the integration process and enabling factors with one caveat. Satisfactory convergence on a solution could only be obtained in the validation sample in the absence of the two item crossloadings accepted in the calibration sample. The same individual factors were identified in the validation sample, judged by their psychometric performance in both single and multiple factor analyses. Fit statistics for both measurement models (and particularly the integration process factor model with higher order factors) were similar in the validation, two-group and calibration sample data sets (see Table 2). RMSEA values in the two-group analysis generally were better.

Standardised solutions for the validation and two-sample (as well as the calibration sample) data sets are reported in Table 4. Most standardised loadings are of comparable magnitude in the validation and calibration samples with the exception of the *Hospital role* factor. This factor was retained, however, on the basis that the two-group analysis that incorporated this factor (see Table 3) produced excellent fit statistics.

These analyses indicate satisfactory confirmation of the re-specified model in the validation sample with only minor differences in model structure and factor loadings. This confirmed model is set out in Fig. 4.

Some of these nine integration process factors and five enabling factors (not requiring re-specification) have been previously described. Those requiring re-specification have not and are set down below.

1. *Holistic patient care* reflects the concept that the GP considers all aspects of the patient including considering family and other relationships.
2. *GP flexibility* reflects the willingness and ability of the GP to work in different and flexible ways in order to accommodate the needs of the patient (it includes use of interpreters and flexible hours).
3. *Patient information* reflects the role of the GP as a guide to the health system and a source of health



(Cross-factor inter-item correlations shown by curved lines (not arrows) to left of model; cross-loading of items on more than one factor are also shown).

Fig. 4. Re-specified and confirmed model of GP integration.

information. It relates to how the GP integrates the patient into the system (all three derived from originally hypothesised factor *Integration with patients* (*Holistic and flexible care*)).

4. *Teamwork* is largely an attitudinal concept and relates to the GP's willingness to work with others and to accept roles as part of a team.
5. *Liaison* reflects the extent of activity communicating on behalf of patients (both derived from originally hypothesised factor *Integration with local providers* (*Teamwork*)).
6. *Care coordination* reflects the concept of the GP as a controller or director of the utilisation of health services and relates to the gatekeeping role GPs have in the Australian health system.
7. *Hospitals* relates to the continuity of GP involvement with their patients throughout an episode of care that involves hospitalisation (both derived from originally hypothesised factor Integration with hospitals and specialists *Care coordination*).

The *Knowledge and education* factor lost the content related to a GP's continuing clinical education and was renamed '*Knowledge of local services*' in recognition of its more restricted content.

The development and validation of an index of GP integration

Scoring the index

Each factor was scored from 1 to 5, these scores being derived by averaging the scores on the five-point Likert scale of the constituent items. This adjustment was necessary as factors varied in the number of constituent items and some missing data existed. Higher order factors were scored by averaging the scores of the contributing scales rather than averaging the scores of all items within the scales. This was done to avoid weighting the score in favour of scales with more items.

In principle item and factor weights are desirable. However, items and factors were not weighted, since at this stage of index development these weights may reflect idiosyncratic sample characteristics. Weights will be developed when data from large, pooled data sets are available.

Reliability and validity

Reliability. Three approaches are broadly used to assess whether test scores are free from measurement error. One is internal consistency—items comprising the factor are assumed to be measuring the same phenomenon. The second is through comparing strictly equivalent forms of the measure. The third is test–retest reliability (repeatability). Pedhazur and Pedhazur Schmelkin (1991) argue that an internal consistency approach is usually superior in the social sciences because truly equivalent indices do not often exist and test–retest reliability is usually confounded by either true change of the measure and/or memory of previous responses at retest.

This is true for this study. There is no equivalent form of the measure and it is difficult to conceive how one could be created. It would require an index to be completed by someone other than the GP (patient or other practitioner in close contact with a GP). It could be argued though, that this person's knowledge of the GP's integration activities is unavoidably incomplete. It is not possible therefore to measure GP self-report bias. If this did exist it can be assumed to be constant and to cancel out when comparisons across time or place are made.

For these various reasons, an internal consistency approach was used instead. The most widely used index of internal consistency is Cronbach's alpha. It has been shown, however, to underestimate reliability substantially where scales are congeneric but not strictly Tau equivalent, as is the case here (Lord & Novick, 1968; Raykov, 1987a, b). Fleishman and Benson (1987) propose an 'index of stability' that involves adjusting the formula for Cronbach's alpha by incorporating weights based on the loadings of each item on the error free score of the scale. The criterion for acceptable levels for the index of stability is the same as for alpha and is commonly set at 0.7. This is though an indicative level and ultimately depends on how much error the user is prepared to tolerate (Pedhazur & Pedhazur Schmelkin, 1991).

Index of stability scores for both the lower and higher order factors were greater than, mostly substantially greater than 0.7 with the exception of the *Liaison* and *Information technology* factors (see Table 3). Index of stability scores for the two higher order factors were *Patient care integration factor*—0.84 and 0.83 and *Public health integration factor*—0.84 and 0.75 in the calibration and validation sample, respectively. These scores indicate acceptable reliability for the index.

Validity. There are limited opportunities to test criterion validity due to the absence of an absolute criterion (gold standard) measuring GP integration. Our focus is rather on construct and content validity. Pedhazur and Pedhazur Schmelkin (1991) identify three steps in construct validation (which also encompass the dimension of content validity as traditionally defined):

1. logical analysis (consideration of the content and logical structure of the construct—critical thinking, theory, knowledge of measurement, design, and analysis are brought to bear).
2. internal structural analysis (evidence that a set of items co-vary as the consequence of a single real construct).
3. cross-structural analysis (evidence supporting hypothesised relationships between the construct and other constructs).

McArdle (1996) uses the terms *internal* and *external validity* in a way that conforms to the second and third of these three steps.

The sequence followed in the development of the Integration index has been designed to comprehensively address the first two of these—logical analysis and internal structural analysis of the GP integration construct. These have been: development of an initial typology; specification of a model of GP integration through the application of Concept Mapping; and re-specification and confirmation of this model by confirmatory factor analysis using SEM based on the two GP national sample data sets.

Cross-structural analysis (tests of hypothesised relationships between the GP integration construct and other constructs) is an activity associated with the (future) application, not the development of the index. The paucity of established theory about the correlates and consequences of GP integration will limit this process. It is also possible to test hypotheses concerning the relationship between the construct and observed variables rather than other constructs. For example, using the calibration data, male GPs scored higher than female GPs on most factors (scales). This did not persist when the association was controlled for hours of work (there was a direct increasing relationship between hours worked and most integration scales). Similarly, GPs who also worked in settings outside general practice, reported higher levels of integration than those who did not. While reassuring, these findings are suggestive rather than conclusive in regard to construct validation.

It can be concluded that construct validity to this point is well supported based on both logic and internal structure analysis though not as yet cross-structural analysis due largely to a lack of established theory.

Discussion

To what extent then has the aim of the study been realised? What is the content and what are the boundaries of the complex, multi-dimensional social concept *GP integration with the health care system*. Does it have a stable reality and describe a stable pattern of events? In the words of the emergent realists, is 'integration' real and not merely a convenient shorthand term to describe a class of programs (Cronbach, 1989)?

The results firstly indicated that GP integration could be conceptualised as integration process (more precisely as two related higher order concepts of patient care and public health integration) as well as integration structure. In other words, three (somewhat less complex) multi-dimensional social concepts were revealed rather than the one complex, multi-dimensional social concept initially explored. The content of the nine factors constituting the integration process model and the five factors constituting the enabling factor model could also be described.

The integration process model with its focus on activity was consistent with the Activity System model of Engestrom (1998). But structure is important too as demonstrated by the existence of the integration structure model. To make this point more specifically, GP characteristics were shown to be important as part of the integration process model though not exclusively so, as shown by the existence of the integration structure model.

Some evidence was provided that these indices were valid, though this still needs to be confirmed in cross-structural analysis investigating hypothesised relationships between the construct and other constructs. In addition, some evidence was also provided that the concepts were real since factor invariance was demonstrated in the validation sample. Again it is still necessary to confirm this in samples drawn from populations at other times and in other places (Marsh, 1993; Meredith, 1993).[1]

Only in this way will it become clear if GP integration is a 'trait'—a stable characteristic of the individual—or 'state'—a characteristic at one point in time and in one set of circumstances. If the latter, it may well be inappropriate to apply an instrument to two populations or at two points in time and conclude that change in a characteristic measured by an instrument (such as GP integration) has occurred. In such a contingency, it would be more appropriate to measure GP integration qualitatively rather than quantitatively (Cunninham, 1991).

In the absence of alternative measures of GP integration and given the importance of integration and GP integration with health system reforms internationally, it is important that these issues be resolved. The indices should be trialled in a variety of contexts in order to better establish whether factor invariance exists.[2] This will necessarily involve defining criteria for factor invariance so that it is possible to distinguish between departures that are significant and non-significant. This is necessary as it must be supposed that some departure (some different factors or items) in different times and places could occur. Trialling should also occur in order to investigate its utility in cross-structural analysis. Given the lack of theory and established evidence about the nature GP integration, it will be unclear to some extent whether results that are not expected (on an atheoretical, 'common sense' basis) are due to weaknesses in the index or to misconceived expectations.

Having said that, given that the index demonstrated factor invariance in the validation sample in a contemporary, national Australian sample, it should be regarded as being of immediate use in studying GP integration in other samples drawn from the population of contemporary Australia. For example, it should first be able to generate, if not test, hypotheses about GP integration and second, compare levels of GP integration between regional groupings of GPs as well as different sub-groups of GPs using these samples.

Acknowledgements

We thank all GPs, consumer representatives, allied health professionals, nurses, specialist doctors, hospital

[1] This is also necessary in order to generate index and factor weights other than the unitary ones used in this paper.

[2] While the response rates in the GP survey were satisfactory judged by other Australian GP surveys, limited information is available on non-responding GPs. This means that factor invariance desirably should be established for this group of non-responding GPs.

administrators and associated professionals participating in Concept Mapping sessions. We also thank all GPs completing the national mail questionnaire surveys. The members of the GP Steering Group comprised A/Profs. Phillip Hegarty, Remi Guibert, Drs. Peter Waxman, Michael Montalto, Stephen Wales and Mr. Richard Smith. The members of the project's Reference Group comprised Drs. Rob Pegram, Peter Joseph, Mary Surveyor, Robert Hall, Neil Beaton, John Aloizos, Charles Block, Chris Peterson, A/Prof. Phillip Hegarty, Prof. Len Gray, Mr./Ms. Shirley Freeman, Demos Krouskos, Jenny Grosvenor, Merinda Northrop. We finally thank Ms. Nancy Carabella who provided administrative support for the project. The project was funded through a General Practice Evaluation Program (GPEP) Program Grant of the Commonwealth Department of Health and Aged Care.

References

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychology*, *37*, 122–147.

Bentler, P. M. (1995). *EQS structural equations program manual*. New York: Multivariate Software Inc., Harper & Row.

Bentler, P. M., & Wu, E. J. C. (1995). *EQS 5.4 for Windows*. Encino. CA: Multivariate Software Inc..

Commonwealth Department of Health & Community Services (1995). *Meeting people's needs better*. A discussion paper Department of Health & Community Services, Canberra.

Cronbach, L.J. (1989). Construct validation after thirty years. In Linn, R.L. (Ed.), *Intelligence: Measurement, theory, and public policy*: *Proceedings of a symposium in honor of Lloyd G. Humphreys*. Champaign, IL: University of Illinois Press.

Cunninham, W. R. (1991). Issues in factorial invariance. In L. M. Collins, & J. L. Horn (Eds.), *Best methods for the analysis of change: Recent advances, unanswered questions, future directions*. Champaign, IL: American Psychological Association.

Dixon, J., Holland, P., & May, N. (1998). Primary care: core values developing primary care: Gate keeping, commissioning, and managed care. *British Medical Journal*, *317*, 125–128.

Engestrom, Y. (1998). The cultural–historical activity theory. http://www.helsinki.fi/~jengestr/activity/6b0.htm.

Everett, B. (1980). *Cluster analysis* (2nd ed.). New York: Halsted Press.

Falik, M., Lipson, D., & Lewis-Idema, D., et al. (1993). Case management for special populations: Moving beyond categorical distinctions. *Journal of Case Management*, *2*(2), 39–45.

Fleishman, J., & Benson, J. (1987). Using LISREL to evaluate measurement models and scale reliability. *Educational and Psychological Measurement*, *47*(4), 925–939.

Galvin, P. F. (1989). Concept mapping for planning and evaluation of a Big Brother/Big Sister program. *Evaluation Program and Planning*, *12*, 53–58.

General Practice Changing the Future Through Partnerships (1998). *Report of the general practice strategy review group*, Commonwealth of Australia.

Harris, M. F., Fisher, R. R., & Knowlden, S. M. (1993). Improving general practitioner involvement in urban hospitals. Departments or divisions of general practice. *Medical Journal of Australia*, *158*, 304–307.

Jöreskog, K. G., & Sörbom, D. (1988). *PRELIS: A program for multivariate data screening and data summarization: A preprocessor for LISREL*. Chicago: Scientific Software Inc..

Jöreskog, K. G., & Sörbom, D. (1989). *LISREL 7*. Chicago: Scientific Software Inc..

Kendrick, T., & Hilton, S. (1997). Primary care: opportunities and threats broader teamwork in primary care. *British Medical Journal*, *314*, 672–679.

Liaw, T., Lawrence, M., & Rendell, J. (1996). The effect of a computer-generated patient-held medical record summary and/or a written personal health record on patients attitudes, knowledge and behaviour concerning health promotion. *Family Practice*, *13*, 289–293.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Canberra: Addison-Wesley.

Macquarie Dictionary (1997). 3rd ed., Macquarie University, NSW.

Markham, K. M., Mintzes, J. J., & Jones, G. M. (1994). The concept map as a research and evaluation tool: Further evidence of validity. *Journal of Research Science and Teaching*, *31*(1), 91–101.

Marquart, J. M. (1989). A pattern matching approach to assess the construct validity of an evaluation instrument. *Evaluation Program and Planning,*, *12*(1), 37–43.

Marsh, H. W. (1993). The multidimensional structure of academic self-concept: Invariance over gender and age. *American Education Research Journal*, *30*(4), 841–860.

McArdle, J. J. (1996). Current directions in structural factor analysis. *Current Directory of Psychological Science*, *5*, 11–18.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika.*, *58*(4), 525–543.

Miles, M. R., & Huberman, A. M. (1988). Drawing valid meaning from qualitative data: Toward a shared craft. In D. M. Fetterman, *Qualitative approaches to evaluation in education*, Vol. 238. New York: Praeger.

Montalto, M., & Dunt, D. R. (1993). Delivery of traditional hospital services to patients at home. *Medical Journal of Australia*, *159*, 263–265.

Øvretveit, J. (1993). *Coordinating community care: Multidisciplinary teams and care management*. Buckingham: Open University Press.

Pedhazur, E. J., & Pedhazur Schmelkin, L. (1991). *Measurement, design and analysis: An integrated approach.*. Hillsdale: Lawrence Erlbaum Associates, Publishers.

Raykov, T. (1997a). Estimation of composite reliability for congeneric measures. *Applied Psychological Measures*, *21*(2), 173–184.

Raykov, T. (1997b). Scale reliability, Cronbach's coefficient alpha, and violations of essential tau-equivalence with fixed congeneric components. *Multivariate Behaviour and Research*, *32*(4), 329–353.

Sharma, S., Durvasula, S., & Dillon, W. R. (1989). Some results of the behaviour of alternative covariance structure estimation procedures in the presence of non-normal data. *Journal of Market Research*, *26*(2), 214–221.

Shumaker, R. E., & Lomax, R. G. (1996). *A beginner's guide to structural equation modelling*. New Jersey: Lawrence Erlbaum Associates.

Southern, D., Batterham, R., Appleby, N., & Young, D., et al. (1999). The concept mapping method: An alternative to focus group inquiry in general practice. *Australian Family Physician-Reason*, *28*(Suppl 1), 35–40.

Trochim, W. (1989a). An introduction to concept mapping for planning and evaluation. *Evaluation Program and Planning*, *12*, 1–16.

Trochim, W. (1989b). Outcome pattern matching and program theory. *Evaluation Program and Planning*, *12*, 355–366.

Young, D., Dunt, D., Southern, D., Batterham, R., Appleby, N., & Liaw, T. (1999). *Integrating GPs with the wider health care system: Development of the integration index for general practice*. Melbourne: Final report to the Commonwealth Department of Health and Aged Care, Department of General Practice and Public Health.