Do growth monitoring and promotion programs motivate mothers to change behaviour?
A systematic review using a combination of Cochrane-style systematic reviews and realist reviews

Research Project in Health Program Evaluation (505-905)
Submitted for the partial fulfilment of
Master of Public Health
School of Population Health
The University of Melbourne

By
Sigit Sulistyo
Student Identification 273749
June 2008
# Table of Content

Disclaimer........................................................................................................................................ iii  
Acknowledgements .............................................................................................................................. iv  
Affirmation ........................................................................................................................................ v  
Abbreviations and Acronyms ............................................................................................................. vi  
Abstract........................................................................................................................................... vii  

Chapter 1: Introduction ....................................................................................................................... 1  
   1.1. Background and Policy Context ............................................................................................. 1  
   1.2. Rationale for conducting this review ....................................................................................... 2  
   1.3. Aims ......................................................................................................................................... 3  
   1.4. Definitions ............................................................................................................................. 3  
   1.5. Structure of this report .......................................................................................................... 4  

Chapter 2: Growth Monitoring and Promotion .................................................................................. 6  
   2.1. Growth monitoring and promotion programs in developing countries ................................. 6  
   2.2. Implementation of GMP in Indonesia ..................................................................................... 7  

Chapter 3: Methodology .................................................................................................................... 10  
   3.1. Why use the realist review approach? ..................................................................................... 10  
   3.2. Why use the Cochrane-style systematic review? .................................................................. 12  
   3.3. Why combine the realist review and the Cochrane-style systematic review? ..................... 13  
   3.4. Outline of method ................................................................................................................ 13  

Chapter 4: Identifying the Review Focus .......................................................................................... 16  
   4.1. Search to provide overview of the GMP program logic ....................................................... 16  
   4.2. Mapping the available evidence ............................................................................................ 18  
   4.3. Specifying the review question .............................................................................................. 18  

Chapter 5: Theory of GMP to Change Mothers’ Behaviour ............................................................... 20  
   5.1. GMP Mechanisms ................................................................................................................ 20  
   5.2. The tailoring intervention and the Health Belief Model ....................................................... 21  

Chapter 6: Searching and Quality Appraisal ..................................................................................... 23  
   6.1. Non-systematic searching ..................................................................................................... 23  
   6.2. Systematic searching .......................................................................................................... 23  
   6.3. Purposive searching ............................................................................................................. 24
List of Figures

Figure 1. The systematic review template. ................................................................. 15
Figure 3. A 3 x 3 matrix of HBM and tailoring intervention strategies .................... 22
Figure 4. Diagram of literature searching process .................................................... 26

List of Tables

Table 1. Comparison between intervention and control groups ............................ 28
Table 2. Study characteristics included in the meta-analysis ................................. 34
Table 3. Summary of primary studies included in the review ................................. 47
Table 4. Contextual constraints identified in the review ......................................... 49
Table 5. Potential corrective actions based on contextual constraints identified ........ 56
Disclaimer

The work in this project was undertaken to partially fulfil the requirements of The University of Melbourne for the degree of Master of Public Health. The views expressed are those of the author and may not reflect the views of The University of Melbourne, School of Population Health.
Acknowledgements
Affirmation

I certify to the best of my knowledge that this report is based on my own research and that I have acknowledge all the work done by others where appropriate, in accordance with the rules and regulations pertaining to academic work at the University of Melbourne. I declare that this work has not been previously submitted for assessment at this or any other institutions.

Sigit Sulistyo

16th June 2008
## Abbreviations and Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMO</td>
<td>Context-Mechanism-Outcome</td>
</tr>
<tr>
<td>GMP</td>
<td>Growth Monitoring and Promotion</td>
</tr>
<tr>
<td>HBM</td>
<td>Health Belief Model</td>
</tr>
<tr>
<td>IMMP</td>
<td>Indonesian Management of Malnutrition Program</td>
</tr>
<tr>
<td>MOH</td>
<td>Ministry of Health Indonesia</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomised-Controlled Trial</td>
</tr>
<tr>
<td>SE</td>
<td>Self-efficacy</td>
</tr>
<tr>
<td>UNICEF</td>
<td>United Nations Children’s Fund</td>
</tr>
<tr>
<td>UPGK</td>
<td>Nutrition Improvement Program</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
</tbody>
</table>
Abstract

Background
The failure of GMP for not achieving its intended outcomes is in part due to the lack of implementation quality. A systematic review on how to maximise the potential of GMP services to improve child nutrition status is needed. The question is to go beyond whether GMP programs work (or do not work), but on how and why GMP programs work (or do not work). There is, however, no single agreed method for synthesising diverse forms of evidence on social interventions such as GMP. This thesis argues that combining Cochrane-style systematic reviews and realist reviews to synthesise the evidence in GMP is feasible and able to produce credible information for decision makers.

Objectives
This paper aims to explore how a systematic review combining both Cochrane-style systematic reviews and realist reviews can help explain one aspect of the GMP program theory and to use this information to inform the evaluation process of the Indonesian Management Malnutrition Program

Methods
To combine Cochrane-style review and the realist review, the realist review is first used to identify the review focus. The Cochrane approach then is applied to guide the searching process for high quality evidence. After finishing the quality appraisal, the realist approach will be applied in the data extraction and synthesis stages.

Main Results
There is weak evidence that use GMP as a tool to personalise risk information alone motivates mothers’ to change behaviour in improving caring practices. Incorporating multiple theoretical concepts, including HBM, tailoring intervention strategies, self-efficacy, and role modelling enhance the likelihood of behaviour change. In addition, this review identified contextual factors that need to be taken into account in applying the mechanisms in different settings. It is clear that the incorporation of multiple theoretical concepts combined with addressing the contextual constraints is important to gain positive outcomes. The nature of these results will be useful to policy makers who need to know not merely whether GMP programs work but these supplies practical theories on what needs to be attended to in constructing better GMP programs.
Conclusion
The review demonstrates that combining Cochrane-style reviews and realist reviews is feasible and helpful as an alternative method to synthesise a complex social intervention where only few evidence on the related topic available such as GMP. The review emphasises the importance of incorporating multiple theoretical concepts and of addressing contextual constraints to maximise the implementation of the GMP activities.
Chapter 1: Introduction

This chapter explores the policy context and rationale of conducting a systematic review on growth monitoring and promotion (GMP) programs. In addition, this chapter explains the objectives of this review, clarifies its definitions, and outlines the structure of this report.

1.1. Background and Policy Context

Growth monitoring and promotion (GMP) is a nutrition intervention that measures and charts the weight of children and uses this information to counsel parents so that they take actions to improve child growth (WHO, 1986). In developing countries, growth monitoring of infants and young children is recognised as a key strategy for prevention of child malnutrition and is an intrinsic part of primary health care.

The rationale for GMP is persuasive, however, there is ongoing debate on the effectiveness of these programs to improve child growth. This debate has currently become heated around the conflicting findings of the evaluation of the Bangladesh Integrated Nutrition Project which incorporated a GMP program (White et al, 2006). This situation stimulates a call for a general review of the impact of large-scale GMP programs to determine if the investments are justified (Panpanich & Garner, 1999).

Responding to this call, Ashworth, Shrimpton, and Jamil (2008) recently undertook a review of the impact of GMP programs in developing countries. They concluded that GMP might not be the best use of limited resources in countries with insufficient health budgets. The authors suggest abandoning GMP activities in settings where this program is not in place or where this program exists but the coverage is low. For settings where GMP programs currently exist (e.g. Indonesia), however, they suggested to maximise their potential rather than stop these programs. This argument is in accord with Coulibaly, Delisle, and Haddad’s (2002) opinion. They proposed that given the evidence that the quality of GMP is low, rather than to discontinue GMP activities, it might be essential first to improve the quality of GMP services.
In view of these opinions, a systematic review on how to maximise the potential of GMP services to improve child nutrition status, particularly in countries where the program already exists, is needed. The question is to go beyond whether GMP programs work (or do not work), but on how and why GMP programs work (or do not work).

Currently, there is no single agreed method for synthesising diverse forms of evidence on social interventions such as GMP. It has been recognised that standard systematic reviews, which normally use the Cochrane-style reviews, focus too much on descriptive causation (what works) and too little on the development of explanations as to why and how the interventions work (or do not work). Recognising the limitations of the Cochrane-style reviews to synthesise evidence on implementation (as opposed to effectiveness), Pawson (2002) proposed a realist review as an alternative method to mitigate these problems. This approach has been criticised, however, for its methodological weakness.

It is therefore reasonable to combine the Cochrane-style systematic reviews and the realist review to profit from their advantages and minimise their weaknesses. This thesis argues that combining the Cochrane-style systematic reviews and the realist review to synthesise the evidence in GMP is feasible and able to produce credible information for decision makers. Further discussion of why this combination is chosen will be discussed in chapter 3.

1.2. Rationale for conducting this review
There are three reasons for conducting a systematic review of the GMP program. First, it has been recognised that the failure of GMP for not achieving its intended outcomes is in part due to the lack of implementation quality. This review will focus on explanatory rather than judgemental synthesis might help to provide credible information of how to improve the implementation of GMP. Putting into Indonesian context, the Ministry of Health Indonesia is currently implementing the Indonesian Management of Malnutrition Program (IMMP), a national community-based nutrition program which incorporates GMP as one of its principal strategies. There is a policy need to review the available
evidence on how to maximise the implementation of the GMP activities integrated in the IMMP.

Second, explicit program theory of the IMMP is not readily available. The concept of program theory refers to the rationale for why and how a program is believed to achieve its intended outcomes. In recent years, there has been an increasing interest in using program theories to assist in designing and conducting an evaluation process in order to provide a firmer basis for program monitoring and impact evaluation (Owen, 2006). Since the IMMP impact evaluation is planned to be conducted in 2009, this current review, which focuses on exploratory analysis, would assist constructing the IMMP program theory. This program theory in turn can be used to guide planning of the impact evaluation for this program.

Third, currently there is a global effort to establish a program theory for GMP programs. This effort is based on the concept that a strong program theory can serve as the foundation for constructing effective programs to improve child growth. This study may contribute to the development of a program theory for GMP.

1.3. Aims
The purpose of this thesis is two-fold:
- To explore how a systematic review combining both Cochrane-style systematic reviews and realist reviews can help explain one aspect of the GMP program theory.
- To use the above information to inform the evaluation process of the Indonesian Management Malnutrition Program

1.4. Definitions
For the sake of clarity, this thesis begins by distinguishing the terms ‘review’ and ‘synthesis’ since they are often used confusingly. For this paper, Mays et al. (2005) definitions on review and synthesis are used to differentiate between these two terms. Review is defined as “the whole process of bringing together a body of evidence which can be drawn from research and other sources, relevant to a particular decision in a policy
or management context” (p. S116). The term synthesis is used to refer to “the stage of a review when the evidence extracted from the individual sources is brought together in some way” (p. S116).

The terms ‘program logic’ and ‘theory’ tend to be used interchangeably. For the purpose of this review, the terms of ‘program logic’ and ‘theory’ are differentiated. The term ‘theory’ in this review is used to refer to ‘systematically organized knowledge applicable in a relatively wide variety of circumstances devised to analyse, predict, or otherwise explain the nature or behavior of a specified set of phenomena that could be used as the basis for action’ (p. 315) (van Ryn & Heaney, 1992). Program logic refers to the chain of causal assumptions that link program resources and activities to accomplish the program intended goals. The theory is often included in the program logic to make explicit the causal assumptions about how an intervention is supposed to work. Thus, program logic is a visual representation of the theory which links programs resources and activities to accomplish program intended goals. The terms ‘theory’, ‘mechanism’, and ‘program theory’ are used interchangeably in this paper. Throughout the paper, the terms ‘carers’ and ‘mothers’ are used interchangeably.

1.5. Structure of this report
Chapter 2 of this report provides an overview of GMP programs in developing countries and summaries the implementation of GMP program in Indonesia. Chapter 3 sets out the method of combining Cochrane-style systematic reviews and realist reviews for searching, prioritising, analysing, and synthesising the literature that was relevant to this review, and gives the research strategy and synthesis methods. Chapter 4 describes the process of formulating the review question and mapping the GMP program theory. Chapter 5 articulates the theory of why growth monitoring is believed to motivate mothers to adopt healthy behaviour. The two theories identified in this chapter presented in a 3 x 3 matrix used to identify underlying mechanisms of the program evaluated in the primary studies during data extraction process. Chapter 6 chapter describes the iterative process in conducting literature searching for evidence. The process employed three strategies in searching literature: non-systematic, systematic, and purposive. Chapter 7 describes each of the primary studies included in the review briefly and interrogates its
results in the light of the theoretical framework established in the chapter 5. Data extraction and data synthesis were produced here using a narrative approach.

The discussion section (Chapter 8) includes three sections. The first part draws together the results of the empirical studies in a particular contexts and circumstances. The second part, it considers the process of undertaking synthesis of complex evidence using a combination of Cochrane-style reviews and realist review. The reviewer reflects on how well this approach work and the challenges that remain. The third part explores the limitations of the current review. Chapter 9 presents the recommendations to inform the evaluation process of the Indonesian Management and Malnutrition Program.

The reviewer has also included two appendices: Appendix 1 and 2 show the critical appraisal checklist for different research designs. Appendix 3 reproduces the data extraction sheet for primary studies included in the review.
Chapter 2: Growth Monitoring and Promotion

This chapter discusses an overview of growth monitoring and promotion (GMP) programs to tackle children nutrition problems in developing countries, including a look at GMP history. In addition, the implementation of GMP program in Indonesia is discussed.

2.1. Growth monitoring and promotion programs in developing countries

In developing countries, GMP activities are an integral part of primary health care. Health workers monitor growth to assess nutrition adequacy and take action in children with evidence of malnutrition (WHO, 1986). Health workers and mothers spend considerable time and resources to conduct this activity (Gerein, 1988).

The use of growth indicators is based on the notion that health and nutritional status of children is best defined by assessing growth measurement (WHO, 1986). There are many techniques of monitoring child growth including weighing, measuring arm circumference, and measuring height. The method of growth monitoring most commonly used is the weighing and plotting weight on growth charts, since weight gain is believed to be the most sensitive indicator of growth (WHO, 1995).

Growth monitoring has a long history. In developing countries, growth monitoring (the first term for GMP before the term ‘promotion’ was added) has been used since 1961, particularly pioneered by Dr. David Morley at Ilesha, Nigeria (Morley, 1962). In the 1970s growth monitoring was implemented in several developing countries in Africa, Asia, Latin America, and the Near East. This program rapidly flourished in the 1980s and gained support from many international agencies including the World Health Organisation (WHO) and the United Nations Children’s Fund (UNICEF) (Ashworth et al., 2008). It has gained popularity in the last three decades and has been practiced in over 154 countries (De Onis, Wijnhoven, & Onyango, 2004).
Growth promotion had always been advocated as an integral part of growth monitoring activities, however, most of growth monitoring activities focused more on weighing and charting and less on promotion (Gerein, 1988). Some authors than call growth monitoring a ‘ritual’ since weighing and charting is done without any actions taken based on the findings (Gerein, 1988). Growth monitoring without any action to take appropriate responses will not be effective in improving child nutrition status (Griffiths, Dickin, & Favin, 1996). The term growth monitoring and promotion (GMP) was therefore introduced to make growth promotion an explicit component (WHO, 1986).

Over the last decade, there has been ongoing debate of the effectiveness of growth monitoring and promotion (GMP) programs to improve child growth. This debate has become heated currently around the conflicting evaluation findings of the effectiveness of the Bangladesh Integrated Nutrition Project which incorporated GMP and food supplementation programs (White et al, 2006). Critics argue that the impact of GMP in various settings is questionable (Chopra & Sanders, 1997; George, Latham, Abel, Ethirajan, & Frongillo Jr, 1993; Gerein & Ross, 1991). Criticisms have been particularly based on the infeasibility of implementing GMP under real conditions. This includes lack of adequate coverage, low mothers’ comprehension on growth charts, inadequate knowledge and skills of health workers in delivering GMP activities, and over emphasis on growth monitoring rather than promotion (Gerein, 1988). On the other hand, GMP supporters believe that by improving the quality of GMP services, particularly by providing appropriate training and supervision of health workers, the available evidence is sufficient to support the continued practice and implementation of GMP (Coulibaly et al., 2002; Ruel, 1995).

2.2. Implementation of GMP in Indonesia

In Indonesia, a GMP program has been implemented since 1970s. The program is part of the Family Nutrition Improvement Program (UPGK) which is a multi-sectoral health program including immunisation, family planning, maternal and child health services (Griffiths et al., 1996). The UPGK is a community-based program which is conducted in integrated health service posts (posyandu) run by village health workers (kader).
The GMP sessions are conducted by *kaders* every month at *posyandu*. They weigh children under five, plot their weight on a growth monitoring card (*kartu menuju sehat*), and conduct nutrition education activities. *Kaders* are trained and supervised by local health officers. The basic objectives of GMP are to detect growth faltering earlier and to change mothers’ behaviour, with the ultimate goal of improved child nutrition through better feeding practices. The GMP program is also used as a tool for community mobilisation by using aggregate weight gain information at the community level to influence community leaders and society to take effective actions in tackling nutrition problems in their community.

Since the economic monetary crisis of 1997, the UPGK seems not to have been as well implemented as previously. This lack of quality of implementation was believed to be one cause of the increasing prevalence of child malnutrition in Indonesia (Ministry of Health Indonesia, 2005). In response to this problem, the Government of Indonesia Ministry of Health has implemented the Indonesian Management of Malnutrition Program (IMMP) (2005-2009). This program aims particularly to revitalise the GMP program that exists but in a dormant condition. The IMMP is four years old and still continues.

The program has three components: (1) training of community health volunteers (*kaders*); (2) growth monitoring and promotion; and (3) supplementary feeding. The key objectives are to improve the caring practices of individuals, households, and communities, and to improve the availability of good quality food for children under five (MOH, 2005).

The program is implemented in all districts in Indonesia. All carers are asked to bring their young children (0-59 months) to *posyandu* monthly for GMP. At each visit, the child is weighed, and the carer receives counselling on health, family planning, breastfeeding, caring practices, and personal hygiene. Based on growth monitoring, children who are categorised as moderately or severely malnutrition children are eligible
for a supplementary feeding program. Carers of children who are receiving food are required to visit the *posyandu* every week to receive the supplement.
Chapter 3: Methodology

The idea of combining the realist review and the Cochran-style systematic reviews was inspired by the work of van der Knaap et al. (2008) of combining Campbell standards and the realist review. The Campbell standards are another type of method for conducting a systematic review that are often used to synthesise evidence of social interventions. In this thesis, the Cochrane-style methods were chosen rather than the Campbell methods since the Cochrane-style methods have a long tradition in reviewing research on the effectiveness of healthcare interventions. In addition, most of systematic reviews conducted in the GMP field used the Cochrane-style approaches (Ben-Joseph, Dowshen, & Izenberg, 2007; Panpanich & Garner, 1999; Roberfroid, Pelto, & Kolsteren, 2007).

3.1. Why use the realist review approach?
Currently, there is no single agreed method for synthesising diverse forms of evidence of social interventions. Mays et al. (2005) argues that the choice of the form of synthesis methods is likely to be crucial related to the review question and the nature of the available evidence. In view of this notion, the realist review is considered the appropriate method to address the current review question and to mitigate the problem of the relative paucity of evidence in the GMP literature.

The review question: Why and how rather than what works
As described in Chapter 2, this review focuses on the question of ‘why’ and ‘how’ GMP programs work (or do not work) rather than on the question of ‘whether’ GMP programs work (or do not work). It has been recognised that the Cochrane-style systematic reviews focus too much on descriptive causation (what works) and too little on the development of explanations as to why and how the interventions work (or do not work); that is why it faces methodological difficulties if being used to answer explanatory questions. For example, Roberfroid et al. (2007) conducted a Cochrane-style systematic review of 22 quantitative studies to assess carers’ comprehension of growth charts. They concluded that “many carers have poor comprehension of growth charts, but their comprehension increases not only with literacy, but also with training” (p. 1074). They, however, did not provide further information about what contextual factors need to be attended to in
enhancing the effectiveness of the training to improve carers’ understanding on growth charts. As consequence, the review findings may have limited use to inform decision-making.

The realist review provides explanatory method to mitigate this limitation. This method is based on the emerging ‘realist’ approach that focuses on Mechanisms, Contexts, and Outcomes—the CMO model. It is argued that the effectiveness of a program is generated through the complex interactions between the mechanisms (M) supposed to underlie it and the contexts (C) in which the program is operating. The O in the CMO model concerns the outcomes of a program, that is, the intended and unintended consequences of interventions that occur from the operation of one or more mechanisms in different contexts (Pawson, 2002).

The realist review offers a model for synthesising evidence that is compatible with the complexities of GMP programs. This model provides an explanatory analysis by identifying the relationship that occurs between mechanisms and contexts in a program that influences its outcomes. The realist review depends on the identification and testing of these mechanisms and contexts to gain an insight into how an intervention works and the extent to which the results can be applied to other settings. It recognises the fact that social interventions exist in open systems with differences in the circumstances; therefore, it never gets implemented in the same manner and never has the same impact (Pawson, 2002).

**The nature of evidence**

It has been recognised that conducting a systematic review on GMP faces difficulties of finding sufficient evidence in this field. For example, Panpanich et al. (1999) conducted a Cochrane-style systematic review to evaluate whether GMP programs work. They excluded all but two trials because their design quality was too low for the results to be valid. They concluded: “There is insufficient reliable information to be confident whether growth monitoring is of benefit to child health” (p 1). In addition, Roberfroid et al. (2007) conducted a systematic review on mothers’ comprehension of growth charts revealed that
there was no study in their database to answer specific questions related to GMP, such as whether GMP is necessary for mothers’ behavioural change. These findings clearly indicate that the available evidence with respect to GMP area is scarce.

The realist review can mitigate this problem since the locus of comparison is not programs but program theories. The unit of comparison of a realist review is on causal mechanisms or program theories that underlying a particular program as to why the program works (or not) for particular groups in particular contexts. The realist review, therefore, gathers evidence not limited only from the GMP field but also from a broad range of knowledge as long as it employs the same program theory.

3.2. Why use the Cochrane-style systematic review?
Popay et al. (2006) argued that the trustworthiness of a systematic review depends heavily on the quality of the evidence base it is built on. The good quality of primary studies and transparency of including the evidence in the synthesis stage therefore are important to ensure the trustworthiness of the synthesis. This thesis is in agreement with this view that internal validation and transparency are important in order to infer credible conclusion in the systematic review.

The realist review, however, has been criticised for its methodological weaknesses regarding the low internal validity and the lack of transparency. It seeks evidence in many forms ranging from newspaper content, personal opinions, opinion surveys, and program guidelines. Many critics on the realist review particularly focus on the lack of quality appraisal protocol in this method, therefore, judge the realist review to be of low internal validity (Dixon-Woods, Agarwal, Jones, Young, & Sutton, 2005). In addition, this method has been criticised for lack of transparency regarding choice of evidence to synthesise (Dixon-Woods et al., 2005). Therefore, it is difficult for other researchers to replicate the synthesis processes to test the credibility of the conclusion it is derived from.

These two limitations, on the other hand, can be mitigated by incorporating the Cochrane-style systematic reviews in the realist review since the Cochrane-style reviews place emphasise on the importance of assessing the quality of research evidence
(Jackson, 2003). Moreover, this method provides transparency by articulating details of how analysis is done and how conclusion is reached (Jackson, 2003). In doing so, the Cochrane-style reviews employ strict inclusion criteria based on relevance and rigour for the primary studies before they enter the synthesis stage (as opposed to the realist review which it appraises primary study in the synthesis phase).

3.3. Why combine the realist review and the Cochrane-style systematic review?
Since the Cochrane-style systematic reviews and the realist review have their own weaknesses and advantages, it is therefore reasonable to combine these two methods to profit form their advantages and minimise their weaknesses. The expectation is that the realist review will profit from the methodological advantages in the Cochrane-style reviews. Accordingly, this thesis argues that incorporating the Cochrane reviews’ principles into the realist review will increase the realist review’s internal validity and transparency.

3.4. Outline of method
There is broad agreement about the main elements involved in undertaking a systematic review. Six main elements are used here, including: identifying the review focus; developing a program theory; identifying studies to include in the review; quality appraisal; extracting the data; synthesising the data; and disseminating the findings (Figure 1). In this template, the stages of the review process are presented as if they occur in a linear fashion simply to provide structure for the clarity of the review process. The review, however, is undertaken more iteratively. For instance, literature searching and mapping the evidence are likely to take place in parallel with finalising the specific review question.

To combine the Cochrane-style reviews and the realist review, the first step is using the realist review approach to identify the review focus. The first task in this step is mapping the GMP program logic, and next, basing on this program logic, the available evidence will be mapped to identify the knowledge gaps in the program logic (Chapter 4). Finally, formulating the mechanisms to be tested (Chapter 5). Once the mechanisms under review have been identified, the next step is searching for potential studies.
The Cochrane approach then is applied to guide the searching process for high quality evidence (Chapter 6). At this stage, criteria are formulated for inclusion and exclusion of study relevance and methodological quality. The search strategies for potential studies are explicitly described including key words, sources of electronic databases, and bibliographies of other studies. Once potential relevant studies have been identified, they are screened for eligibility according to the inclusion and exclusion criteria. After selecting the relevant studies, the quality of these studies is determined. Preferably, this process is done by more than one researcher, which in turn raises the need for strategies to resolve disagreement. This requirement, however, could not be satisfied in the current review since this is part of academic assessment which should demonstrate an individual work.

After finishing the quality appraisal, the realist approach is applied in the rest of the review stages. The next step concerns the production of data extracts (Chapter 7). At this stage, the available studies are summarised using a tabulation form. These studies are extracted based on the realist configuration of Mechanisms-Contexts-Outcomes. After finishing the data extraction, all the studies are synthesised to reconcile the different outcomes between studies by exploring relationships between the mechanisms and the contexts (Chapter 7).
Figure 1. The systematic review template. Adapted from Pawson (2006).

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identifying the review focus</td>
<td>Searching for primary studies</td>
<td>Quality appraisal</td>
<td>Extracting the data</td>
<td>Synthesising the data</td>
<td>Disseminating the findings</td>
</tr>
<tr>
<td>Map key program theories</td>
<td>Background familiarisation search</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Map the available evidence</td>
<td>Search for available evidence testing the program theory</td>
<td></td>
<td>Annotation, note taking</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specify the review question</td>
<td>Search for sources of program theory</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Formalise mechanisms to be tested</td>
<td>Search for empirical studies to test model</td>
<td>Assessment of relevance and rigour</td>
<td>Collation of materials from selected primary studies</td>
<td>Absorbing primary materials into developing synthesis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Search for further empirical studies on revisions to model</td>
<td>Further assessment of rigour as each study enters the synthesis</td>
<td>Detailed reportage evidence from each case study</td>
<td>Juxtaposing, adjudicating, reconciling, consolidating, and situating further evidence</td>
<td>Theory re-articulation</td>
</tr>
</tbody>
</table>
Chapter 4: Identifying the Review Focus

This chapter describes the first step of the review process of identifying the review focus. This step is important since getting the review focus ‘right’ is critical to the success of the systematic review process overall. The initial step was constructing the program logic of GMP. The program logic then was used to map the available evidence and in turn, the knowledge gaps that need to be addressed by a systematic review were identified. Based on the program logic and the initial mapping of the available evidence, the scope of the review was refined.

4.1. Search to provide overview of the GMP program logic
Program logic describes the chain of causal assumptions that link programs resources, activities, intermediate outcomes, and ultimate goal. Developing a program logic at an early stage of a review can provide a valuable guideline to map the available evidence and to inform decision about the review question. In addition, this exercise helps to articulate theories or mechanisms that are plausible of how and why a GMP program is believed to achieve its intended outcomes.

Although a simple technology of weighing and charting is at the center of GMP in developing countries, GMP is a complex operational communication intervention. Various steps in the process need to be addressed for a correct implementation. A wide variety of information sources was used in surfacing and articulating the program logic of GMP including procedural GMP guideline published by the World Bank, primary studies on GMP, and official bulletins. The program logic of GMP obtained from these documents formulated in an if-then diagram was summarised in Figure 1.
Figure 2. The GMP program logic
The available systematic reviews on certain components of the program logic are shown in grey cycles.
The above constitutes the basic program logic of GMP as implicitly is depicted in the GMP policy guidelines and evaluation studies on GMP. Basically, the GMP program theory consists of two main arms. First, GMP is used as a screening tool to early detect growth faltering in children. Second, GMP is used as a motivational tool to deliver health messages in order to motivate mothers to adopt healthy behaviour. In theory, for a GMP program to achieve its goals, every step in the sequence needs to be in place as a failure at any one step may discontinue the causal linkages and thus compromise the intended outcomes.

4.2. Mapping the available evidence
This paper is not intended to be a comprehensive review of the abundant literature on GMP given the resource constraints. Therefore, the review needs to be focused on certain aspect of the GMP program theory. One of the considerations of choosing what aspect of GMP program theory wants to be reviewed is based on the available systematic reviews have been conducted and the remaining knowledge gaps that can be filled. This exercise would avoid unnecessary duplication of effort.

Some systematic reviews have been conducted to test some aspects of the GMP program logic (depicted in grey cycles in the figure 1). There are two systematic reviews addressing a question whether mothers have a sufficient comprehension on growth charts (Ben-Joseph et al., 2007; Roberfroid et al., 2007). There is also a small body of systematic reviews on some of the issues related to the effectiveness of GMP as a screening tool to early detect growth faltering in children (Roberfroid, Kolsteren, Hoeree, & Maire, 2005).

4.3. Specifying the review question
Using the program logic to guide the evidence mapping, it is now clear the knowledge gaps in the program logic that have not been evaluated systematically (depicted in the figure 1 as dashed frames), these including the ability of staff on delivering GMP services and their understanding on GMP philosophy. In addition, there are some knowledge gaps in the communication component of GMP such as the use of GMP as a tool to motivate behaviour change. The current reviewer was interested to elaborate more on this aspect.
The communication component of the GMP program logic posits two mechanisms for its effectiveness: to help mothers understand concepts of normal and abnormal growth; and to motivate behaviour change. Robrefroid et al. (2007) have synthesised the evidence regarding the first mechanism, but there is no systematic review in relation to the second mechanism. Little is known of the additional impact on behavioural change obtained by the inclusion of GMP in a nutrition education intervention. This review seeks to unpack the mechanisms of how a GMP program works to motivate mothers to improve caring practices (or why it fails) in particular contexts.
Chapter 5: Theory of GMP to Change Mothers’ Behaviour

The next task is to articulate the theory of why GMP is believed to motivate mothers to adopt healthy behaviour. The information particularly comes from documents and procedural guideline on the implementation of GMP from the World Bank (Griffiths et al., 1996) and the World Health Organisation (WHO, 1986). The World Bank and WHO are the leading institutions who proposed the use of GMP to tackle malnutrition in developing countries.

5.1. GMP Mechanisms

There are at least two mechanisms underlying the assumption that GMP is believed to motivate mothers’ behaviour change. The first mechanism is using GMP as a tool to increase mothers’ perceived risk on their children nutrition/health status. GMP uses this principle: the child’s growth curve drawn on the chart and compared to the reference curve to show the mother whether her child is at risk, and provides prompt response of her efforts to help the child. This mechanism is based on the assumption that a prerequisite for behaviour change is an individual recognising that he or she is at risk. For this mechanism to be effective, the use of growth charts is important since small changes in a child’s growth may not be easily detected simply by observing the child (Griffiths et al., 1996).

The second mechanism is using growth monitoring as a tool to individualise education messages (Griffiths et al., 1996; WHO, 1986). When nutrition education is tailored to growth monitoring results, it is more specific and relevant for mothers to understand than when it is offered as general messages. By counselling the mother in this tailored way, there is more likelihood that she will do what is specifically suggested than if she attended a group lecture and had to select by herself the most relevant advice for her child (Griffiths, 1988).

These two mechanisms, not surprisingly, have been used in other health programs to enhance behavioural change. The first mechanism is consistent with the perceived risk, which is one component of the Health Belief Model (HBM). The second mechanism is consistent of the principles of tailoring intervention.
5.2. The tailoring intervention and the Health Belief Model
Tailored interventions are an area of research in behaviour change that is developing rapidly. In the literature, tailoring of interventions has been used in public health programs since the 1990s. Hawkins et al. (2008) have suggested a definition for tailoring as:

Any of number of methods for creating communications individualized for their receivers, with the expectation that this individualization will lead to larger intended effects of these communications. (p. 1)

The mechanism of how tailored intervention works is still in development. One of the tailored theory frameworks was proposed by Hawkins et al. (2008). This study provided a valuable insight for the current review’s effort to test the mechanism of tailoring intervention embedded in the GMP. This study constructed a theoretical framework of the mechanisms of why and how tailoring enhances health communication. The authors identified three distinct strategies for tailoring to achieve its goal in behavioural change, namely personalisation, feedback, and content matching.

Personalisation strategies aim to promote attention by expressing that the communication is designed specifically for the receiver. These tactics are intended to make behavioural information more relevant and meaningful to the receiver. Feedback attempts to report back to individuals’ information about themselves which is collected from individual assessment or elsewhere. These strategies aim not only to increase attention, but also focus on key determinants of health behaviours. Content matching attempts to assess key individuals’ behavioural determinants and to use this information to formulise individual messages.

In practice, tailored interventions are used in combination with other health behaviour theories such as in combination with the HBM. The HBM posits that the decision of whether an individual adopts a healthy behaviour is a function of individuals’ perceived risk of a negative health outcome and the perceived benefits minus barriers to take action of healthy behaviours (Janz & Becker, 1984). From this perspective,
when an individual recognising that he or she is at risk, the relative balance of benefits and barriers determines the likelihood of protective action occurring. Those who are viewing more benefits than barriers are more likely to take action than those viewing more barriers than benefits (Janz & Becker, 1984).

The HBM and tailoring intervention are used to guide the review process to identify the underlying mechanisms of a program in the primary studies included in the review. In the literature, these two mechanisms are often used in combinations, therefore, this combination is presented in a 3 x 3 matrix in which the components of these mechanisms intersect (figure 3). This matrix is used to identify underlying mechanisms of a program evaluated in the primary studies during data extraction process (chapter 7). The key mechanisms identified from the primary studies are indicated by the symbol $M_n$.

**Figure 3. A 3 x 3 matrix of HBM and tailoring intervention strategies**

In the matrix above, the rows represent all components of tailoring strategies, whereas the columns represent the HBM components. The grid can be used to locate all combinations in HBM and tailoring strategies used as mechanisms underlying an intervention. Note that this matrix is not intended to be exhaustive since there are some studies which incorporate other mechanisms not captured in this matrix. The point is merely to produce a model capable of dealing with a variety of combinations in using HBM and tailoring intervention strategies in the literature.
Chapter 6: Searching and Quality Appraisal

This chapter describes the iterative process in conducting literature searching for evidence. The process comprised three stages. First, the literature searching was begun with a non-systematic searching of the GMP literature in an attempt to map the GMP program logic. Secondly, a systematic searching of GMP literature was conducted to identify existing studies on GMP programs that could potentially be used to test the mechanisms under review. It was found during this searching process that there were only small number of studies on GMP satisfied the inclusion criteria. Therefore, in the third stage the inclusion criteria were broadened to include broader literature outside the GMP field. Since time limitation, the literature searching in this stage employed a purposive rather than a systematic strategy.

6.1. Non-systematic searching
The non-systematic searching actually began in the early steps of this review, but for the reason of systematic heading, this process is discussed pos-hoc in this chapter. The early non-systematic searching provided much of the background to the study. In addition, this preliminary searching process aimed to construct the GMP program logic. This process retrieved GMP guidelines published by World Bank and WHO, primary studies on GMP, and official bulletins.

6.2. Systematic searching
This section focuses more on searching for primary studies included in the synthesis process. This section is particularly using the Cochrane approach in locating primary studies and in appraising the studies quality to filter high quality studies included in the review.

To obtain the database for the systematic review, a keyword search of electronic databases was conducted. Specifically the reviewer searched MEDLINE, POPLINE, ERIC, and the Cochrane Library with the keywords: ‘growth monitoring’ or ‘growth chart’. The search was limited to the title/abstract field, which yielded 141 references. For pragmatic reasons this review rejected all titles whose full-text paper was not available in English or Indonesian (the languages spoken by the reviewer). Articles retrieved through the
electronic search were screened to assess if they met the inclusion criteria based on relevance and rigour.

The relevance aim to address the question of is the paper relevant to the mechanisms want to be tested, in other words, does the evidence shed light on understanding the mechanism under review (i.e., using a combination of the HBM and the tailoring interventions strategies to motivate behaviour change). It is not about whether the study cover particular topic on GMP. Since the time constraint, the reviewer read only the abstract to decide the relevance of the studies to test the mechanisms. Only studies identified as using mechanisms of at least a combination of one of the HBM components and one of the tailoring strategies were included in the preliminary primary studies in the review. On the basis of relevance, 12 of these items were retrieved.

The next step was to appraise the studies quality. This review only included studies that used randomised or quasi-randomised comparison of GMP programs. The experimental research designs (randomised controlled trials, non-randomised controlled trials) and the quasi-experimental designs were evaluated using modified versions of the quality checklist developed by Khan et al. (2001) (Appendix C). The studies ranks were grouped as ‘outstanding and essential to include’; ‘some limitations and essential to include’ and ‘many important limitations’. This review rejected all studies ranked as ‘many important limitations’. This process resulted in two articles being included in the review. Since only two relevant controlled trials retrieved in this review, the study quality details were presented descriptively along the study synthesis (rather than on the matrix format) to allow readers to interpret the evidence in the light of any limitations.

6.3. Purposive searching
Recall from the methodology section, in the realist synthesis, the unit of analysis is mechanism rather than program. Since only two studies on GMP that met the inclusion criteria, the literature on tailored interventions outside the GMP field was included to gain further explanation of the inner mechanisms of GMP. Including the literature on tailoring interventions provided rich insight to refine and search further explanations of how growth monitoring can be used to enhance health communication.
Since time constraint, the searching strategy used in this step was more purposive than systematic. The aim was to retrieve studies that directly related to the mechanisms under review. For pragmatic reason, the reviewer first attempted to search available meta-analyses in tailoring intervention. The meta-analyses are argued providing the highest evidence since they aggregate from many high quality experimental studies. The reviewer searched in the same database with the keywords: ‘meta-analysis’ And ‘tailor*’ Or ‘individual*’ Or ‘personal*’. The search was limited to the title/abstract field. Articles retrieved through the electronic search were screened to assess if they met the inclusion criteria.

The inclusion criteria were based on relevance and rigour. The relevance is that the extent to which the study can be used to test the mechanisms under review. Only studies identified as using mechanisms of at least a combination of one of the HBM components and one of the tailoring strategies were included in the review. The quality of the meta-analyses was assessed based on quality criteria suggested by Oxman et al. (, 1994 #589) (Appendix C). The studies ranks were grouped as ‘outstanding and essential to include’; ‘some limitations and essential to include’ and ‘many important limitations’. The reviewer rejected all studies ranked as ‘many important limitations’.

The search yielded five meta-analyses, these studies than appraised the extent to which the meta-analyses relevant to the mechanisms under review and the quality of the meta-analyses. There were two meta-analyses which met the inclusion criteria. Because of resource constraints of this current review, only one meta-analysis was retrieved which was considered as the most relevant to the mechanism under review. In addition, six of the 57 primary studies included in the meta-analysis were retrieved since they were centrally relevant to test the program theory. Reasons of why these studies were included in the review were further discussed in the synthesis section (chapter 7). It was assumed that the primary studies included in the meta-analysis were in a high quality.

The diagram of the searching procedure was described in figure 1. In total, there were nine studies which was ultimately included in the synthesis stage.
Figure 4. Diagram of literature searching process

Non-systematic searching

Systematic searching

Purposive searching

Retrieved 6 primary studies that centrally related to test the theory under review
Chapter 7: Data Extraction and Synthesis

The emphasis in this chapter was to refine the theory about the mechanisms and contexts in which combining HBM and tailoring interventions can be effective. This chapter did not attempt to search for best practices, however, it attempted to test further of the theory under review. Each of the studies was described briefly and its results were interrogated in the light of the theoretical framework established in the chapter 5. Data extraction and data synthesis were produced here using a narrative approach. In addition, a tabulation of summary was presented to give an overview of the Contexts-Mechanisms-Outcomes identified in each of the primary studies (appendix 1).

The reviewer found nine key studies which relevant to test the mechanism of why and how GMP is believed to motivate mothers to adopt healthy behaviours. The following studies were selected purposefully, as an optimal set capable of putting to test the theory. Two of these studies related to the field of GMP area. Seven of them came from the broad studies on tailoring interventions. These studies were included since they had explanatory value for the mechanisms under review. Because of resource constraints of this current review, the reviewer did not retrieve primary studies if a high-quality meta-analysis had included them, unless they were centrally relevant to test the program theory.

The realist formula context + mechanism = outcome was used to identify and to code the theories that were believed to influence the process of the program in achieving its intended outcome. The mechanisms and key contextual factors extracted for the purpose of theory refinement were indicated by the symbol Mn (as depicted in the 3x3 matrix) and Cn respectively. The supportive contexts (contexts that produce successful outcomes) were indicated as C+n, whereas the constraint contexts (contexts that induce failure) were indicated as C-n.

Ruel et al. (1991) conducted a pre and post-intervention design study to test whether growth charts increased the impact of nutrition education on maternal learning about weaning practices and diarrhoea in Lesotho. Mothers attending each of nine health centers in Lesotho were sequentially assigned to one of three groups: either to one of two types of growth chart (n = 756) or to no charts (n = 268). There was no difference between the two groups in terms of socioeconomic and demographic characteristics.

The intervention offered was similar in all other respects, that is, monthly sessions of GMP along with individual counselling and group nutrition education. Growth chart teaching was done both in small groups and individually with mothers. The issue covered included, the rationale for using growth chart, the meaning of two axes and the printed lines and the interpretation of individual patterns. For individual teaching, the child’s own growth pattern was discussed with the mother and the direction of the growth curve was interpreted. The authors provided summary of interventions in each groups, reproduced here as table 1.

<table>
<thead>
<tr>
<th>Activities</th>
<th>Visit number</th>
<th>No Chart Group</th>
<th>Chart Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Growth monitoring</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Child weighing</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Plotting of weight on chart</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Interpretation of weight pattern</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Without growth chart</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Individual counselling of mothers</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Growth chart</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Feeding practices</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>
From this table it can be seen that this study was built around the mechanism that a visual depiction of the baby’s growth in the growth charts (opposed to growth monitoring without charts) can be used to increase mothers’ perceived risk (M₁) of their child nutrition status in order to foster their learning of nutrition education. Such mechanism assumed that carers have a good comprehension of growth charts and growth is assessed regularly in order the growth pattern to be meaningful. There is, however, no attempt in this study to ensure that these assumptions were satisfied.

In the light of the 3 x 3 matrix, the current reviewer argues that in this study, the GMP program employs combination of perceived risk and personalisation strategy (M₁+M₂), increasing mothers’ attention on the health/nutrition messages by personalising risk information of their children nutrition status depicted from the growth chart. There was no effort to assess either mothers’ perceive barriers or attempt to match the content of the messages based on child nutrition status since all mothers received the same education messages regardless their children nutritional status.

Maternal knowledge about weaning practice and diarrhoea management was tested at baseline and 4 months later. The baseline knowledge scores were not significantly different between the two groups. At the end of the study, the groups with growth charts scored only slightly higher overall than those without (77% and 75% respectively).

This finding indicates that the use of GMP mainly to personalise risk information (M₁+M₄) was not effective to enhance maternal learning on health/nutrition education. This premature inference is tested in the following studies in this review to test whether this inference also replicate in other studies. This study, however, did not evaluate whether this mechanism has effects on mothers’ behaviour change. Other studies in this review provide further enlightenment to test this mechanism.

Despite there was no randomisation in this study, the authors assured the reader that confounding factors were less likely to contaminate the results. This was evidenced by no difference in socioeconomic and demographic characteristics between the two groups. Moreover, multivariate analysis was also conducted to control for potential confounding
 factors including baseline scores, maternal and child characteristics. Since there was no blinding technique to assign project staff to the study group, project staff might give additional attention and time to the growth chart group (Hawthorn effect). If Hawthorn effect would have happened, the estimated effect of GMP would be an overestimate of its real effect. Thus, it was unlikely affect the evidence being used in this review that the overall impact of using GMP to personalise risk information on fostering learning effect was small.


In the study 1, the authors evaluated effects of growth monitoring using growth charts in fostering maternal learning on education messages. George et al. (1993) evaluate similar intervention but assess the long term effect of whether a GMP program in South Indian Villages improve child’s nutrition status. The research strategy employed was a cluster randomised-controlled trial. In this study, they compare 12 villages which were divided into 6 ‘growth monitoring package of interventions (GMP) villages and 6 ‘non-growth-monitoring package’ of interventions (NGM) villages. The authors assured the reader that growth monitoring was successfully implemented in GMP villages, based on correct used of growth chart and good mothers’ understanding on growth chart.

These villages were paired with respect to caste composition, access to main road, cropping patterns, and distance from rural clinics. The aim of the study was to test if GMP in conjunction with an intensive education intervention produced any additional benefit on weight and height gain compared with education intervention alone. The difference between these two groups was:

“NGM villages received the same interventions as did GMP villages except for growth monitoring, ie, education without charts. The content of educational messages in both groups was identical but it was imparted to mothers in NGM villages without the use of growth charts.” (p. 348)
Based on this study description, the underlying mechanism built to this study, similar to the Ruel et al. study, was using the growth chart as a tool to personalise risk information ($M_1+M_4$) without any efforts to address psychosocial barriers. In this study, the personalised risk communication was promoted with presenting children growth pattern to their mothers, which in turn could increase mothers’ attention of their child health/nutrition status, therefore motivate them to improve caring practices and lead to improve child growth.

The authors found that after 30 months of interventions similar improvements in growth were seen in GMP and NGM children. In concordance with the study by Ruel at al., this study demonstrates that using personalised risk communication ($M_1+M_4$) had no effect on children nutritional status.

In the discussion part, the authors provided further exploration of why the GMP did not achieve its intended outcome. This provides further information to identify the contextual constraints that influence this failure. The authors assured the reader that in the study area, the mother got additional support from the health worker to persuade the decision-making members of her family to take action. They found that the mother’s improved self-esteem coupled with the confidence that she could get support to provide better child care appeared to be gained with or without the use of growth charts. Bear in mind that in this society, women have lower position than their husbands evidenced by they often do feel helpless because of frequent verbal abuse, and not uncommonly physical violence, alcoholism, and or promiscuity of their husbands. This finding indicates that in the context where women have lack of decision-making power ($C_1$), it seems that delivering personalised risk information may not be sufficient to enable mothers to change behaviour. In the context when women have good decision-making power, this mechanism might be able to motivate mothers to improve their caring practices.

Second, the authors concluded that sensitising mothers to the faltering growth of their children (by looking at growth patterns) did not provide better care since budgetary constraints in their family. This demonstrates that personalised risk communication may
not be effective to motivate mothers to change behaviour if they have budgetary constraints ($C_{-2}$).

The authors also identified another contextual constraint of implementing the personalised risk communication strategy. In this study, growth monitoring had not been conducted in the villages before. In the intervention group, growth monitoring was started late because many community members were unwilling to allow their children to be weighed. The introducing growth monitoring in GMP villages, moreover, even resulted in significantly lower weight gains in younger children of GMP villages in intervals of the first year. This may indicate that the cultural inappropriateness ($C_{-3}$) may compromise recipients intention for behaviour change.

In this study, we can find that the GMP was not effective to increase child growth if it was solely used to deliver personalised risk information ($M_{1}+M_{4}$), without any effort to deliver relevant message to address mothers’ barriers to adopt the healthy behaviour (the contextual constraints: $C_{-1}$, $C_{-2}$, and $C_{-3}$). This is relatively safe inference since it gives support to the earlier study (Ruel & Habicht, 1991) about how the use of growth charts without any effort to address psychosocial determinants resulted in its inability to foster maternal learning from nutrition education. This embryonic theory is tested further in the following studies.

The research strategy employed in this study was a cluster randomised-controlled trial. The randomisation was successful in distributing the potential confounding evenly evidenced by similar characteristics in social, demographic, service use, child age distribution and child nutrition status at baseline between these two groups. There was no information whether the project staff was blinded to the study groups that could have introduced bias where project staff give additional attention and time for the GMP groups (Hawthorn effect). If the bias did occur, the estimated effect of GMP would be an overestimate of its real effect. Thus, it was unlikely affect the evidence being used in this review that the overall impact of using GMP to personalise risk information on mothers’ behaviour change was insignificant, at best.

Noar et al. (2007) conducted a meta-analysis of 57 experimental and quasi-experimental studies of tailored print health behaviour change interventions. The reason for including a meta-analysis in the current review is twofold. First, Pawson (2006) argues the rationale of inclusion of meta-analyses in the realist review is “to demonstrate the idea of explanatory refinement” (p. 144) since it operates at high level of aggregation. Second, from the methodological quality point of view, Cochrane methods value a meta-analysis as the highest level of evidence, since it aggregates multiple experimental designs.

The meta-analysis drew on literature through to the end of 2005 and aimed to assess not only whether tailored print interventions work, but also to address the question of why and how they work (or fail) by examining several sets of moderators that might impact the effects of tailoring. The effect sizes from data reported in the article were calculated using the Pearson correlation coefficient $r$. There was evidence that tailoring print health interventions were effective with the sample size-weighted mean effect size of the effects of the intervention being found to be $r = .074$. This net impact measure pooled together data on a diversity of program mechanisms and goals with a wide variety of program contexts. As such, however, the net effect does not provide explanatory value for the current review.

The moderator analysis in the meta-analysis provides more valuable insight since it tried to unpack the relative effect of different types of theoretical concept underpinning tailoring. The theoretical concepts identified including Trans-Theoretical Model, self-efficacy, Stages of Change, Theory of Reasoned Action, Theory of Planned Behaviour, Health Belief Model (HBM), and Social Cognitive Theory. This part of the meta-analysis was particularly important in the current review since it provided statistical estimates of which theoretical concepts were associated with the more positive effects. More pragmatically, it also assisted the current review since the authors had categorised the primary studies included in the meta-analysis based on their underlying theoretical concepts. This allowed the current reviewer to by-pass the step of identifying the theoretical concepts in each of
the 57 studies and to focus on analysing the mechanisms and contexts in the individual studies to test the theory under review.

Interestingly, this study found that the only theoretical concept associated with significantly decreased effect sizes was tailored perceived susceptibility. This inference is consistent with the two previous studies (study 1 and 2) that the use of GMP as a tool to personalise risk information did not have any impact on mothers’ behaviour change.

This statement needs further elaboration to identify the mechanisms and contexts to provide explanatory values in refining the theory under review. In the meta-analysis, the authors identified six studies considered as using HBM as the theoretical concept to tailor the interventions. Since perceived susceptibility is one component of HBM, the current reviewer focuses to careful unpack the individual studies which incorporated HBM as identified by Noar et al. The authors provided summary of these interventions, reproduced here as table 2.

<table>
<thead>
<tr>
<th>Study</th>
<th>Health Behaviour</th>
<th>Follow-up</th>
<th>Comparison</th>
<th>N</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nansel et al. (2002)</td>
<td>Pediatric injury prevention</td>
<td>3 weeks</td>
<td>Message</td>
<td>213</td>
<td>0.231</td>
</tr>
<tr>
<td>Baker et al. (2004)</td>
<td>Flu vaccination</td>
<td>12 months</td>
<td>Message</td>
<td>1,373</td>
<td>0.068</td>
</tr>
<tr>
<td>Campbell et al. (1994)</td>
<td>Keeping pediatric preventive medical appointments</td>
<td>1 week</td>
<td>Message</td>
<td>183</td>
<td>0</td>
</tr>
<tr>
<td>Jibaja-Weiss et al. (2003)</td>
<td>Cervical cancer screening</td>
<td>1 year</td>
<td>Message</td>
<td>984</td>
<td>-0.270</td>
</tr>
<tr>
<td>Kreuter et al. (1996)</td>
<td>Seat belt use</td>
<td>6 months</td>
<td>Message</td>
<td>535</td>
<td>0.016</td>
</tr>
<tr>
<td>McCaul et al. (2002)</td>
<td>Mammography screening</td>
<td>6 months</td>
<td>Control</td>
<td>1,177</td>
<td>0.065</td>
</tr>
</tbody>
</table>

Note. N = sample size; r = sample size-weighted mean effect size

From the above table, it is interesting to ask why the Nansel et al. study was the only study that revealed the highest weight effect, whereas the Jibaja-Weiss et al. study showed
negative weight effect. The meta-analysis, however, did not provide an explanation for this particular association and the current reviewer needs to look at each of the primary studies for the appropriate clues. While the aggregation method used by these authors to evaluate evidence might arguably have been appropriate for their own research question, it is inappropriate for this review research question, to elaborate more about the mechanisms and the contexts that might produce different outcomes.

The reviewer, therefore, retrieved these studies to unpack the mechanisms and contexts of why these interventions achieved different outcomes using narrative approaches. Each of these studies was described below.


This study evaluated the effectiveness of patient-specific letters to improve the show rate of well-child appointments. The study was undertaken in predominantly poor urban children in New York which 71% of visits were made by Medicaid recipients. The RCT compared two groups: letter group and postcard group. Parents in the letter group received patient-specific and visit-specific reminder letters; those of the postcard group received reminders specified only the date and time of the appointment.

The letters were designed utilising components of the HBM. They were the date and time of appointment; described the general risk of not being immunised and provided age-specific information on the benefits of immunisations and screening tests, and guidance on nutrition, safety, development.

These descriptions suggest that the basic mechanism of risk communication ($M_1$) used in the previous two GMP studies also applies in this study. In addition, the researchers incorporated perceived benefit based on age specific ($M_3+M_5$) in the letter group. This study adds a specific exploration for this question: whether incorporating ‘matched benefit information based on age specific’ to risk information ($M_3+M_5+M_1$) adds much to behaviour change.
After 13 months, the researchers found no significant difference among the groups. The authors’ exploration of the reasons for these unexpected results provides further insight about contexts. They speculated that the failure of the letter reminders might result from parents already knowing the messages conveyed (C-4). This explanation was supported by a study in the same population which showed that 96% of parents believed that measles was serious in severity, and 91% felt that immunisation was important to prevent disease (Roghamn, Rodewald, & Szilagyi, 1993), and therefore, they would not benefit from further teaching via the letter. The authors also speculated that the lack of difference between the letter and postcard group might be due to unconvincing messages conveyed by the letters (C-5). Moreover, it is important to note that the intervention in this study is only matching the messages based on age specific, rather than to individual characteristics. Less precise tailoring (C-6) might also affect the ineffectiveness of the letter reminders.

Despite this study targeting different behaviour outcomes from the two previous GMP studies, many of the same contextual constraints apply. Thus, the ineffectiveness of the letter reminder might be because of budgetary constraints (C-2) hinder the parents to immunise their children. This constraint may apply in this study since only 71% visitors were under Medicaid. Other parents may use fee-for-services type of payment which might hinder them to immunise their children. The authors however did not provide further information regarding this issue.

In the same manner, contextual constrains identified here might be applied in GMP programs. Mothers would not change their behaviour if the same messages were always repeated (C-4) in each GMP session. It is also make sense that tailored information addressed specifically on child need (e.g. individual feeding difficulty, food preferences), would be more effective than based on age specific (C-6). The most important one perhaps is the credibility of the messages (C-5), if mothers do not trust the trustworthiness of messages delivered by health workers, they less likely to adopt the messages given to them.
Up to this point, however, it is still not clear whether the null outcomes in the study 1, 2, and 4 were due to wrong mechanisms or due to contextual constraints identified in each of the studies. This embryonic theory, therefore, needs further refinement by exploring following studies in the review.


This study provides further evidence of the inner mechanisms of combining HBM and tailoring interventions. The inquiry compares three types of mailings to persuade women to obtain a mammogram. These include: (1) a simple reminder message; (2) a reminder emphasising personal risk; and (3) a reminder tailored to the participants’ chief barrier to getting mammogram. The simple reminder letter outlined perceived susceptibility by stating ‘Breast cancer remains the number one cause of cancer deaths for women in North Dakota’. It continued by noting that mammography can be effective for early detection and that Medicare helps pay for a screening mammogram.

Second choice, the risk letter included two additions to the reminder letter. First, one sentence in the letter reminded the readers that the risk of breast cancer increases with age. Second, the researchers attached a one-page insert that pictured an older woman from a rural North Dakota town and a headline stating ‘Your risk of breast cancer increases with age’, and the insert ended with the line ‘Ask your doctor about regular mammograms. . . because once is not enough.’

The third option, the postcards first sent to tailored-letter participants to identify what was preventing them from having mammogram. The researchers then sent letters tailored to each woman’s chief barrier based on the information collected in the postcards. The tailored-letters were similar to the reminder letter except for a paragraph addressing the aim barrier the women reported.

There seems an approximate fit with some combination of HBM and tailored interventions here. The basic idea under scrutiny was that by addressing individual barriers, in addition
to general risk information \((M_2+M_5)+M_1\), would foster the screening rate compared to using only risk information \((M_1)\). The background to this intervention was the assumption that risk information alone is insufficient to change complex behaviour without addressing psychosocial barriers. This assumption fits rather neatly with the George et al. hypothesis in the study 2 that noted addressing participants’ barriers might improve the effectiveness of personalised risk information.

The outcome evaluation, surprisingly, showed no net impact. The reminder messages which incorporated personalised risk information in this experiment, whether accompanied by additional material to address individual perceived barriers or not, failed to increase mammography rates.

The authors speculated on an explanation for this seemingly perverse outcome. They suggested that this failure was due to no specific action instruction \((C-\gamma)\) accompanying the remainders. The letter reminders did not articulate a specific action regarding the barriers that preventing the women from having a mammogram. For example, in this study the most frequently reported barrier was no time. For women who said they ‘didn’t have time,’ the letter suggested, “Try scheduling a mammogram around a special date, like your birthday”. The letters would have been more persuasive if they incorporated a specific instruction such as contacting a local clinic to schedule the mammogram that suit with their time and listing the most near clinic locations and operating hours.

The current reviewer elaborated on other possible contexts that may explain this counter-intuitive outcome. A notion from the study 4 is helpful in explaining the null effects of the letter reminders. The message of “Try scheduling a mammogram around a special date, like your birthday” might have been familiar to the recipients \((C-\gamma)\), thus further repetition would not add much in encouraging them to obtain the screening.

Putting this theory in the GMP provides a further nuance for understanding the GMP theory. No specific action instructions \((C-\gamma)\) are real issue in the GMP. If health workers only describe the nutrition status depicted in the growth charts and provide general advice based on mothers’ barriers, would less likely lead to behaviour change. This notion is
further discussed in study 9, of the need for incorporating self-efficacy for mothers to be able to change their behaviour.


This is a randomised controlled trial to compare the effects of different types of computer-generated, mailed reminders on the rate of influenza immunisation. The patients were randomised to one of four interventions: (1) no reminder; (2) a generic postcard; (3) a personalised postcard from their physician; and (4) a personalised letter from their physician, tailored to their health risk.

The standard message of the printed materials was based on the HBM, included a description of who is at risk of contracting influenza, a statement of the fact that influenza can be serious, and assurance that the vaccine is safe and effective. The printed materials also advised individuals to get the influenza vaccine, and listed the influenza clinic locations and operating hours. The generic postcard included only the standard content message. The personalised postcard was from the primary care physician, addressed to the patient at risk, and contained the standard message. The personalised tailored letter was from the primary care physician, addressed to the patient at risk, and contained a message tailored to the patient’s risk factors for influenza.

This is the excerpt of the patient letter:

The flu or pneumonia can strike anyone but it is more serious in people who [merge field – criteria options]. . . . Receiving these shots cannot cause you to get the flu or pneumonia . . . You don’t have to make an appointment . . . The costs of the shots are covered by Medicare and Health Alliance Plan.

From the description of the interventions, the current reviewer argues that the tailored message incorporated HBM and personalisation strategy as its mechanisms. From the letter above, it can be identified that all HBM components were used in the letter included perceived susceptibility ($M_1$), perceived barrier ($M_2$), and perceived benefit ($M_3$).
letter also employed a personalisation strategy \( (M_4) \) since it used patient name and emphasise individual risk factors.

The immunisation rate was measured after 12 months of intervention. The authors found that all three of the remainders studied increased the influenza vaccination rate when compared with the control group, however, the absolute increase appeared small.

The results from this study showed that personalised risk communication combine with provision of messages addressing general perceived barriers and perceived benefits \( ((M_1+M_4)+M_2+M_3) \) favour healthy behaviour change, although the effect was small. It is important to note that in this study, there were no budgetary barriers for the patients to get the vaccination since the vaccination was free and accessible \((C_2)\).

This study provides further refinement of the GMP theory. GMP theory which usually only emphasise on personalising risk information, might profit by incorporating other HBM components \( (M_2+M_3) \). This theory, however, only be effective (even the effect was small) if there is no resource constraint to access the intended behaviour change \((C_2)\). This primitive theory is refined in the subsequent studies.


This is the only study that revealed negative outcomes. Essentially, it is another example of frustration about how a personalised risk information intervention failed to work. The researchers conducted a randomised-controlled trial examined the effectiveness of a personalised form (PF) letter and a personalised tailored (PT) letter in promoting compliance to cancer-screening recommendations. The target population was 1574 urban low-income and minority women in Houston, Texas.

The PF letter contained generic cancer information about risk factors for breast and cervical cancer; and the importance of screening and early detection. Whereas the PT contained minimally tailored individualised risk information based on medical records data
affected breast and cervical cancer screening. The intervention descriptions suggest that the authors compared the effectiveness of personalised risk information ($M_1+M_4$) vs general risk information ($M_1$). The program theory is based on familiar idea that delivering messages relevant with individual perceived risk is efficacious in change behaviour.

One year after the intervention, surprisingly, the PF group was significantly more likely to schedule a screening appointment and to have undergone a Pap test and mammography than were the PT and control groups ($P < 0.001$). In this study, letters that contain tailored individualised cancer risk factor information were less effective than letters that contain generic cancer risk information to foster women receiving cancer screening.

A noteworthy finding is the failure of combination between ($M_1+M_4$) involved in promoting behaviour change. The authors speculated that this counter-intuitive finding might be due to the tailored letter was too alarming for the recipients ($C_8$) which made them anxious about undergoing screening test. They also pointed out the problem of unreliability of medical record data which might be outdated or inaccurate to individualise the message. This suggests that when the recipients do not confident with the credibility of the messages ($C_5$), tailored intervention even can compromise the recipients’ intention to change behaviour.

Putting this theory in the GMP programs gives a further insight for understanding the GMP theory. The alarming messages ($C_4$) are factual issues in the GMP. If the messages delivered by health workers are too alarming, mothers might too anxious to come to the subsequent GMP sessions and surrender for behaviour change. The inaccuracy of message ($C_5$) is also important to be considered in GMP since many health workers had insufficient knowledge and skills on health/nutrition education. If mothers do not trust health workers’ competency in delivering the health/nutrition messages, it is makes sense that they will not follow the health workers suggestions.

This RCT evaluated the effectiveness of interventions using an enhanced Health Risk Assessment (HRA) to change in seven different behaviours. The mechanism under scrutiny, namely HRA, is actually one component of the HBM (i.e. perceived susceptibility). The difference of HRA and HBM is that the risk information provided in typical HRA only influence a user’s perceived susceptibility, but—as opposed to HBM—it has no means of addressing perceived benefits and barriers. Again, this study re-emerges familiar question on mechanisms to be tested: What is the effect of enhancing perceived susceptibility without addressing the barriers and benefits of reducing the risk?

The researchers tested three variations of HRA: enhanced HRA feedback, typical HRA feedback, or HRA without feedback (control). Enhanced HRA feedback consisted of combination HBM and tailoring mechanisms: personalised risk information \((M_1+M_4)\), content matching on barriers \((M_2+M_3)\), content matching on reasons for wanting to change \((M_3+TTS)\), content matching on health benefits \((M_3+M_5)\), self-efficacy (SE), and past attempts and failures to change the behaviour. From the description from this study, it showed that the intervention employed more complex combinations than the previous studies, including \((M_1+M_4)+(M_2+M_3)+(M_3+M_5)+(M_5+TTS)+SE\).

Once again—similar with other previous studies—this study revealed mixed results. After six-month follow-up, participants receiving enhanced HRA feedback were 18% more likely to change at least one risk behaviour than were those receiving typical HRA feedback or no feedback. The enhanced HRA feedback succeeded to encourage participants to adopt healthy behaviour in cholesterol screening, dietary fat consumption, and physical activity, but not in smoking, seat belt use, mammography, and pap smears.

These findings, however, indicate that combining many theoretical concepts does not automatically lead to positive outcomes. It is therefore important to go in-depth to identify the constraint contexts. The researchers noted that the fact that tailored messages were effective in some behaviours but not in others might be related to the number of behaviours...
considered in the letters. In this study, participants received tailored feedback on three or more behaviours. Without prioritisation on certain behaviours to be addressed (C-9), it might lessen the impact of tailored message to affect individual behaviour change.

This study provides a consistent pattern of evidence along with the findings from study 4 and study 5, namely incorporating more complex combinations between HBM and tailoring strategies would increase the likelihood of behaviour change. This impact might be enhanced by focusing on one or two behaviours rather than addressing more than four behaviours.

The contextual constrains identified here might be applied in GMP programs. Mothers would not change their behaviour if excessive information were given to address various behaviours. This is an important issue since mothers usually have low education background to understand and remember all the information given by health workers. The health worker should be able to prioritise one or two behaviours prioritised to be addressed in the first place.


This randomised controlled trial compared the effectiveness of tailored and generic persuasive communication delivered in a primary care setting on the adoption of home and car safety behaviours. During routine well-child visits, a primarily African-American samples of parents of children ages 6-20 months (n = 213) was randomised to receive either tailored or generic information regarding the prevention of injuries to their child.

Participants were asked to complete baseline assessments using a computer. These assessments aimed to measure injury risk behaviours, injury prevention behaviours, and relevant psychosocial constructs. Based on these assessments, tailored messages were constructed using a computer program. The tailored injury prevention handout consisted of three pages that contained:
“... a description of the respondent’s injury scores, along with motivational messages tailored to the respondent’s injury-related locus of control and self-efficacy. ... general information on the importance and relevance of the injury area, specific information about the highest priority behaviour contributing to the child’s risk for injury, and specific steps to take to reduce risk. Each page also contained a role-modelling scenario of a parent describing why and how she or he decided to use the specific safety practice being recommended to the participant. The role modelling scenario was tailored to the individual’s beliefs about the effectiveness of the particular safety practice and the difficulty of implementing the safety practice.”

The above description suggests that the key feature of printed material in the intervention group was combinations of HBM components and tailoring strategies. The combinations include personalised risk information \((M_1 + M_4)\), content matching on barriers \((M_2 + M_5)\), and perceived benefit \((M_3)\). In addition, the authors also incorporated role modelling and self-efficacy. Role modelling is built from the social learning theory, which is related to the capacity to learn by observing both the behaviour of others and the rewards received for different patterns of behaviours (observational learning). The handout also mitigated some of the contextual constraints identified in the previous studies, namely specific instructions \((C+7)\) and prioritisation strategy \((C+9)\).

It is important to note that in this study each of the participants was asked to complete the baseline assessment using computer. The assessment questions took 10-15 minutes for most participants to finish. Compare to the previous studies, this study put considerably effort in assessing ones’ information used to tailor the education messages. The detailed assessment of baseline characteristics \((C+10)\) might lead to produce more credible and valid tailoring messages compare with tailoring using information retrieved from the information depicted in the simple postcard questionnaire (study 5) or from the medical records (study 7).

It can be noticed that many of the mechanisms already discussed are revisited in this particular program. It seems that this program was reflected in attempts to remedy what are
considered weaknesses in earlier prototypes. For instance, the researchers incorporated specific instruction (C7), focused only on two behaviours to be modified (C9), and used detailed baseline assessment to identify individual risk behaviours and psychosocial constructs (C10).

At follow up, participants who received tailored information reported greater adoption of home and car safety behaviours than those receiving generic information. In addition, within the tailored information group, those who discussed the information with their physician showed significantly greater change than those who did not. However, this difference was not observed among those receiving generic information. This suggests that counselling might enhance the effect of tailoring interventions.

A familiar enough pattern is reported here. Findings from this study support the previous inference that more complex combination HBM items with tailored strategies, more effective in promoting health behaviour change than only using incomplete combination. Moreover, addition of other mechanisms such as self-efficacy and role modelling might enhance the positive effect. This inference is safe, since it is in line with the meta-analysis in study 3 which noted that studies which tailored on 4-5 theoretical concepts had significantly larger effect sizes than those tailored on 0-3 concepts. Moreover, this study emphasise the importance of counselling to reinforce the effect of tailoring messages.

It is also important to note that provision of a specific action instruction (C+7) seems important to the successful of this intervention. In particular, this finding gives support to the study 5 that found combination of HBM and tailored tactics was not effective, in part was due to the absence of specific instructions in their mailed messages (C+7).
Chapter 8: Discussion

This chapter provides discussion of the synthesis process and the methodological combination. Again, the configuration of Mechanisms + Contexts = Outcomes is used to structure the first discussion part of this chapter. In addition, strengths and limitations of the current reviews are highlighted.

8.1. Synthesis processes and results
To the best of the current reviewer knowledge, this study is the first systematic review combining Cochrane-style systematic reviews and realist reviews used to synthesise the evidence on GMP programs. A similar methodological approach had been conducted to combine Campbell standards and realist reviews in synthesising the evidence on criminal justice (van der Knaap et al., 2008).

The exercise of developing a program logic as a preliminary step in conducting a systematic review appears to be an effective tool for organising complex ideas about a program. The program logic can help to articulate the underlying mechanisms of the causal chains in the logic model. Moreover, this exercise can help to more accurately identify the knowledge gaps to focus the review questions. This is particularly important since there is evidence of replication of effort in conducting systematic reviews on GMP. For example, the systematic reviews conducted by Robrefroid et al. (2007) and Ben-Joseph et al. (2007) basically addressed similar review questions.

Despite the GMP programs have been implemented in the past three decades, this review only found two high quality studies that assessing communication component of the GMP. Most of published evaluation papers yielded during the literature search concentrated mainly on weighing process, whereas communication component of GMP as an important issue has not been dealt with adequately.

The review processes in the synthesis stage particularly tracked down the program theories (mechanisms and contextual factors) in each of the primary studies and brought together these primitive theories to inform the refinement of the GMP program theory. The aim was
to try to extract and articulate a general theory about the circumstances that either constrain or support the use a combination of HBM and tailoring intervention strategies as a program mechanism to motivate behaviour changes. The combination of HBM and the tailoring intervention strategies (depicted in a 3 x 3 table) provided a useful framework for identifying the underlying mechanisms in each of the primary studies.

Table 3. Summary of primary studies included in the review structured based on the realist configuration of Mechanisms + Contexts = Outcomes

<table>
<thead>
<tr>
<th>Study</th>
<th>Mechanisms</th>
<th>Contexts</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Constraint</td>
<td>Supportive</td>
</tr>
<tr>
<td>1</td>
<td>(M₁+M₄)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>(M₁+M₄)</td>
<td></td>
<td>C₋₁, C₋₂, C₋₃</td>
</tr>
<tr>
<td>3</td>
<td>((M₃+M₅)+M₁)</td>
<td>C₋₄, C₋₅, C₋₆</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>((M₂+M₅)+M₁)</td>
<td>C₋₇</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>((M₁+M₄)+M₂+M₃)</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>6</td>
<td>(M₁+M₄)</td>
<td>C₋₄, C₋₅</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>(M₁+M₄)+(M₂+M₅) + (M₃+M₅)+SE+ (M₅+TTS)</td>
<td>C₋₅</td>
<td>+</td>
</tr>
<tr>
<td>8</td>
<td>(M₁+M₄)+(M₂+M₅) + (M₃+M₅)+SE+Role modelling</td>
<td></td>
<td>C₊₇, C₊₉, C₊₁₀ ++</td>
</tr>
</tbody>
</table>

From the table 3 above, it is evident from this current review that GMP programs not only share common mechanisms with other health interventions, but also that the contextual factors that promoting or constraining their success occur and reoccur across different health domains. While the primary studies relate only on particular behaviour characteristics of child caring practices, immunisation, mammography, and so on, the mechanisms and contexts which associate with success and failure tend to replicate between the studies.

**Mechanisms**

Emphasising personalised risk by using growth monitoring is reasonable because perceived risk is an important motivator of many self-protective health behaviours. This review, however, noted that personalised risk information alone may not sufficient to affect the likelihood of action. This inference closely confirmed with the meta-analysis.
finding on tailored printed material to motivate behaviour change that showed the only theoretical concept associated with significantly decreased effect sizes was tailored perceived susceptibility (Noar et al., 2007). In addition, this conclusion is consistent with theories of behaviour change that suggest risk information alone is seldom sufficient to change complex behaviours (Janz & Becker, 1984).

This inference does not imply that tailored information based on HBM is not effective in change behaviour. Perceived risk is only one component in the HBM and this review suggests that combining perceived risk with other constructs of the HBM (i.e. perceived barrier and benefit) may increase the effectiveness of tailoring communication.

In addition to combining HBM with the tailoring strategies, this review provides evidence that incorporating other theoretical concepts favour for behaviour change. The other theoretical concepts identified in this review including self-efficacy and role modelling. The meta-analysis also gives quantitative confirmation for this inference which showed that combining multiple theoretical concepts would significantly increase the efficacy of a tailoring printed material intervention (Noar et al., 2007).

Self-efficacy and role modelling are of important in the GMP practices since there is a growing attempt to incorporate positive-deviant approach, which is incorporating the self-efficacy and role modelling, in conjunction with the existing GMP programs (Olga, Keeley, Burkhalter, & Bashir, 1997). Positive-deviant approach is a method employing mothers in the local community whose children are well-nourished but come from the lower economic half of the community as a role model. This approach uses a combination of HBM, role modelling, and self-efficacy. It recognises the importance of addressing barriers to access nourished food by discovering affordable and nourishing local foods that mothers can give to their children. This information is based upon information gathered in their own community from mothers of well-nourished children from the lower economic half of the community—these are the positive-deviant mothers. A mother also has the opportunity for observational learning (role modelling) as she watches the positive-deviant mothers model the new feeding and caring practices. In addition, this approach emphasises the importance of belief in mothers’ ability to successfully perform a healthy behaviour.
(self-efficacy) by incorporating local practices and wisdoms to convince mothers that they could have well-nourished children.

The incorporation of a positive deviant approach in coordination with a GMP program has been proven to be feasible and more effective than provision only a GMP program (Olga et al., 1997). This review provides support to the growing evidence of the importance of this combination. However, much work remains before it can be clearly understood how these different components interact and contribute to program success.

**Contexts**

This review identified contextual factors that need to be taken into account in applying the mechanisms in different settings. These contexts were summarised in Table 4.

**Table 4. Contextual constraints identified in the review**

<table>
<thead>
<tr>
<th>Contextual constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>(C₁) Lack of decision-making power</td>
</tr>
<tr>
<td>(C₂) Budgetary and time constraints</td>
</tr>
<tr>
<td>(C₃) Cultural inappropriateness</td>
</tr>
<tr>
<td>(C₄) Repetitive of known messages</td>
</tr>
<tr>
<td>(C₅) Lack credibility of messages</td>
</tr>
<tr>
<td>(C₆) Less precision in tailoring</td>
</tr>
<tr>
<td>(C₇) No specific action instructions</td>
</tr>
<tr>
<td>(C₈) Too alarming messages</td>
</tr>
<tr>
<td>(C₉) No prioritisation</td>
</tr>
<tr>
<td>(C₁₀) Less detailed assessment on baseline characteristics</td>
</tr>
</tbody>
</table>

As this list is not intended to be exhaustive, or that it captures precisely the key contextual limitations on combining HBM and tailoring intervention strategies to motivate behaviour change, therefore it should be seen as preliminary findings. Nevertheless, the reviewer believes that the nature of these results will be useful to policy makers who need to know not merely whether GMP programs work (or do not work) but these supplies practical theories on what needs to be attended to in implementing GMP programs to achieve better outcomes.
Outcomes

It can be argued that variation of outcomes identified in the primary studies might be influenced by the types of behaviour being addressed. For example, risk perception are probably more important for behaviours that it is easy to carry out (e.g., immunisation) than when it is difficult to carry out behaviour (e.g., mammography). It is also recognised that the variation of outcomes in the studies included in this review might be influenced by the variation of sample sizes used in each of the primary studies. The meta-analysis, however, provides quantitative confirmation that these potential confounders are not significant in influencing the outcomes of the primary studies retrieved from the meta-analysis (Noar et al., 2007).

Overall, the primary studies included in the current review shows mixed-findings of the effectiveness of combining the HBM and the tailoring intervention strategies. It is clear, however, that the incorporation of all key components in the HBM and the tailoring intervention strategies combined with addressing the contextual constraints is important to gain positive outcomes. This pattern can be seen in the table 4.

While these findings add to the growing body of evidence to evaluate the GMP program theory, less is known on health workers capability to implement these theories into practice. The reviewer argues that the program theory identified in this review is perhaps feasible to be enforced in the field setting. As health workers performance on understanding growth charts and counselling can be enhanced by appropriate training (Roberfroid et al., 2007); therefore, it might be feasible to retrain health workers to be able to implement this program theory in practice. In addition, some nutrition education studies in developing countries using nutrition counselling delivered by health workers revealed that nutrition education that addressing behaviour determinants (include perceived barrier and self-efficacy) could be effective to improve children growth (Bhandari et al., 2004; Penny et al., 2005; Santos et al., 2001).

Although this review focused only on the communication aspect of GMP programs, these findings also can be attributed to other interventions that incorporate similar mechanisms.
For example, efforts to increase immunisation rates might benefit from employing combination of HBM components and tailoring strategies taking into account the contextual factors identified in this review.

8.2. Methodology Combination
The current review demonstrated that the combination of the Cochrane-style systematic reviews and the realist reviews providing useful information for decision makers of how and why the GMP can be used to motivate mothers’ behaviour change. The usefulness of combining these two methods also has been demonstrated by a recent realist review on efficacy of school feeding programs. Greenhalgh et al. (2007) analysed 18 studies of school feeding programs in disadvantaged children included in a Cochrane systematic review using the methods of a realist review. They came up with identifying factors that explaining the success or failure of the programs that illuminate the findings from the Cochrane review.

Perhaps the most important feature of this current review was its ability to provide credible evidence in the situation where there is a relative paucity of evidence in the GMP area. This issue has been recognised as the major obstacle of conducting a systematic review related with GMP topics (Panpanich & Garner, 1999).

The reviewer recognised shortcomings of combining these two approaches. First, most of the RCTs and quasi-experimental studies included in the database provided little information on comprehensive program theories. This finding conforms the observations made by van der Knaap et al. (2008) that found very few studies in their review articulating an explicit program theory. This review therefore supports the needs for future studies on complex intervention such as GMP programs to incorporate explicit program theories. This notion is in accord with the UK Medical Research Council’s recommendations on the design and evaluation of complex health interventions (Campbell et al., 2000).

The second drawbacks of combining these approaches is that in this current review, most of the studies give minimal information with regard to contexts underlying the
interventions being evaluated. This issue is partly because most of the studies did not document sufficient detail descriptions of the unique aspects of different implementation in their respective contexts and settings. Such lack of detail description hinder the reviewer of identifying contextual factors underlying the success or fail of mechanisms incorporated in each of the programs evaluated in the primary studies.

8.3. Study limitations
Limitations of the study should be noted for the reader to use the review findings. The appraisal for relevance and study quality was based only one reviewer which could possibly increase the likelihood of bias. The Cochrane-style reviews suggest including more than one person in conducting a systematic review. This requirement is impossible to be satisfied in the current review since this project is part of the assessment to complete an academic degree which has to demonstrate an individual performance. The reviewer, however, attempted to minimise this potential bias by using an explicit theoretical framework to assess studies relevance and by using standard quality appraisal checklists for the quality appraisal.

Another limitation of this review is the small number of studies included in the synthesis. Popay et al. (2006) argued that the trustworthiness of a systematic review depends on the quality and quantity of the evidence included in the review. This study satisfied the first criterion by including only high quality primary studies based on the Cochrane methods. The reviewer, however, only retrieved eight trials included in the current review consists of two trials on GMP programs and six trials retrieved from the meta-analysis on tailoring printed materials. Nevertheless, this number is sufficient to provide illumination of the limited number of available studies on GMP related to the review question. The analysis, however, would be undoubtedly be enriched by inclusion of other relevant studies in tailoring intervention areas, such as interventions using a tailoring telephone counselling which have grown rapidly (Champion et al., 2007).

It is important to note that seven of nine studies included in this review were on tailored printed material. It has been recognised that tailored printed-messages may not be persuasive enough compare to tailored counselling. A study by Champion et al. (2007)
showed that tailored counselling was more effective than tailored printed material to promote uptake in mammography program. This indicates that GMP might even be benefited from incorporating the theory refined in this review since GMP program theory is using a combination of tailored printed material (the growth charts) and tailored counselling (by health workers). Further studies incorporating tailored counselling would add significant evidence to the current review. Nevertheless, despite the fact that all of the primary studies on tailoring interventions included in this review were limited to tailored printed material (rather than emphasise on tailored counselling), many of the mechanisms and contextual factors apply to GMP program theory.
Chapter 9: Conclusion and Recommendations

9.1. Conclusion
The thesis has showed how the Cochrane-style systematic reviews and the realist review can be combined to provide credible information on the GMP programs. The methodological approach adopted has enabled this thesis to mitigate the major obstacle of absence of evidence in GMP area. In addition, this combination might improve the value of the systematic review in informing decision-making process.

This review provides some evidence of the failure of using GMP only to personalise risk information to affect behaviour change. Moreover, this study shows that the incorporation of other theoretical concepts combined with addressing the contextual constraints is consistent with better outcomes. It seems reasonable, therefore, to conclude that combining the HBM components, tailoring intervention strategies, self-efficacy, and role modelling, incorporated with addressing the contextual constraints would be important for the successful of the GMP implementation.

Despite the reviewer search for further details about the mechanisms and contexts related to GMP, the reviewer considers that the quantity of trials retrieved were insufficient to provide a comprehensive evidence based for policy makers. Nevertheless, the review provides some important practical insights for policy makers and practitioners to maximise the implementation of GMP, particularly in Indonesia.

9.2. Recommendations
The recommendations are discussed in three sections: future research, the IMMP formative evaluation, and the IMMP summative evaluation. The recommendations for the IMMP formative evaluation was structured based on the mechanisms and contexts identified in the review.

9.2.1. Future research
There is already a small body of research which tested some of the causal linkages in the GMP program logic as described in the Chapter 3. For the remaining knowledge gaps
identified in the program logic, future systematic reviews need to fill these gaps. These include: (1) health workers’ knowledge and skills in delivering the GMP activities; (2) health professionals’ understanding and perceptions of GMP philosophy; and (3) role of GMP in encouraging health workers to deliver effective counselling. Filling these gaps will contribute to the refinement of the GMP program theory.

9.2.2. The IMMP formative evaluation

Potential corrective actions based on the mechanisms identified

It is evidenced in the current review that incorporating more complex theoretical concepts would lead to behaviour actions. The existing GMP programs in Indonesia, and in many developing countries, tend to focus on using of GMP as a tool to deliver personalised risk information without take into account the need to address mothers’ barriers and benefits. This review emphasised the need to incorporate tailoring strategies in addressing recipients’ barriers and benefits.

In view from this notion, it is recognised the fact that a national GMP program never gets implemented identically and never has the same impact, because of differences in the barriers in each setting where the GMP programs are implemented. General guidelines would be hard to apply where there would no doubt be a vast variation in mothers’ social and cultural barriers. It is important, therefore, to identify potential local barriers for behaviour change. This research should focus on identifying the community-specific attitudes, perception, and barriers of child-care in general and feeding practices in particular. Results from such research are important for tailoring the general education messages into context specific information.

This review also shows the benefit of incorporating additional mechanisms of self-efficacy and role modelling, incorporated in the positive deviance approach, to the combination of HBM and tailoring strategies. GMP programs in Indonesia might benefit from incorporating the positive deviant approach along side the GMP activities, which has been proved more effective in motivating mothers to improve caring practices than only GMP alone.
Potential corrective actions based on the constraint contexts identified

The potential corrective actions based on the contextual constraints identified in the review were summarised in the table 5.

Table 5. Potential corrective actions based on contextual constraints identified

<table>
<thead>
<tr>
<th>Contexts</th>
<th>Constraint(s) identified</th>
<th>Potential corrective actions</th>
</tr>
</thead>
</table>
| (C₁) Lack of decision-making power | In Indonesia, the authority of older family members and male is dominant. | • Strengthening mother’s skills in problem-solving and decision-making.  
• Providing specific coping strategies for dealing with situations where women have lack of power so they can anticipate and be ready to counteract the anxiety aroused by this situation |
| (C₂) Budgetary and time constraints | In some parts of Indonesia, mothers do not have access to resources to provide good caring practices. | • Incorporating income generation activities in the IMMP  
• Improving women’s access to resources they need to maximise their contribution to family nutrition. |
| (C₃) Cultural inappropriateness | The local food preparations are various across different parts of Indonesia | Incorporating local knowledge and experience in food production, care, and feeding practices. |
| (C₄) Repetitive of known messages | Health workers often provides similar messages and repeat in each GMP sessions | Encouraging health workers to listen to assess mothers’ baseline knowledge and their learning needs to ensure that they deliver messages what mothers do not know, or reinforce what they do know. |
| (C₅) Lack credibility of messages | | Providing easy flow chart material to equip health workers with hands on references in delivering health/nutrition education messages. |
| (C₆) Less precision in tailoring | • Lack of supervision and training leads to low credibility of messages delivered by health workers  
• Health workers tend to provide general advice than specific instructions | |
<p>| (C₇) No specific action instructions | | |
| (C₈) Too alarming messages | Health workers often deliver too alarming messages regarding nutrition status depicted on the growth charts | Improving health workers ability to deliver appropriate messages and treat the mothers with respect to create a safe and relax environment to learn. |</p>
<table>
<thead>
<tr>
<th>Contexts</th>
<th>Constraint(s) identified</th>
<th>Potential corrective actions</th>
</tr>
</thead>
</table>
| (C₉) No prioritisation | Messages tend to be vague and too many messages delivered at the same time (e.g., oral rehydration solution, immunisation, family planning, hygiene, etc) | • Focusing on one or two topics and messages based on magnitude of the behaviours.  
• Supervisors have to monitor accuracy and clarity of messages and adherence to prioritisation concept. |
| (C₁₀) Less detailed assessment on baseline characteristics | Lack of time often hinders health workers to obtain detailed assessment of psychosocial determinants that prevent behaviour change | Allocating adequate time allowed for nutrition counselling. |

**Health worker training**

From the potential corrective actions identified above, it is no doubt a need to provide adequate trainings of health workers, especially in equipping and enabling them to provide effective counselling and mitigating the contextual constraints identified in this review. The stakeholders, therefore, need to put sufficient investment in behaviour change communication training for health workers.

**9.2.3. The IMMP Summative evaluation**

- This review shows that only few evaluations of the GMP programs which take into account communication component of GMP to motivate behaviour change. This indicates that most of the evaluations did not consider the theory underlying GMP program in guiding the evaluation. It is therefore a need to incorporate theory-based evaluation in evaluating GMP programs to explore explicitly linkages between every important activity that leads to intended outcomes in guiding the evaluation process.

- The IMMP evaluation should avoid questions structured only to judge causal relationships, such as ‘Does the IMMP program increase nutrition level of children enrolled in the program?’ Rather, evaluation questions should also be framed with explanatory view, for example, ‘What contextual factors do account for the success (or failure) of the IMMP program?’
• The IMMP summative evaluation should document comprehensively distinctive aspects of the IMMP implementation particularly on particular contexts and settings. Such detailed descriptions can be used to inform future program development and allow for comparison across settings. It is suggested, therefore, for the IMMP evaluation to incorporate qualitative research methods along side the usual quantitative methods. This mixed method might be a particularly useful method since it will enable researchers to consider effectiveness (outcomes) and implementation (mechanisms and contexts) of the program being evaluated.
References


Roghmann, K., Rodewald, L., & Szilagyi, P. (1993). *Study to increase immunization coverage in inner-city areas*: Center for Disease Control and Prevention


Appendices

Appendix 1. Quality checklist for experimental (randomised and non-randomised controlled trial) and quasi-experimental designs. Adapted from Khan et al. (2001)

1. Was there a clear research question, and was this important and sensible?
2. Was the assignment to the treatment groups really random?
   If the study was non-randomised, could a randomised design have been used?
3. Was the treatment allocation concealed?
4. Were the groups similar at baseline in terms of prognostic factors?
5. Were the eligibility criteria specified?
6. Were outcome assessors blinded to the treatment allocation?
7. Was the care provider blinded?
8. Was the patient blinded?
9. Were the point estimates and measure of variability presented for the primary outcome measure?
10. Did the analysis include an intention to treat analysis?

Global rating for this checklist:
1. Outstanding and essential to include (more than eight items answered Yes)
2. Some limitations and essential to include (five to seven items answered Yes)
3. Many important limitations (less than four items answered Yes)
Appendix 2. Quality checklist for meta-analysis studies. Adapted from Oxman et al. (1994)

1. Did the meta-analysis address focused question?
2. Were the criteria used to select articles for inclusion appropriate?
3. Is it unlikely that important, relevant studies were missed?
4. Was the validity of the included studies appraised?
5. Were assessments of studies reproducible?
6. Were the results similar from study to study?

Global rating for this criteria:
1. Outstanding and essential to include (six items answered Yes)
2. Some limitations and essential to include (four to five items answered Yes)
3. Many important limitations (less than three items answered Yes)
### Appendix 3. Tabulation of primary studies.

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Health behaviour</th>
<th>Mechanisms</th>
<th>Contexts</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ruel et al. (1993) Lesotho</td>
<td>Cluster trial before and after</td>
<td>Knowledge in weaning and diarrhoea management</td>
<td>(M₁+M₄)</td>
<td>Generic messages</td>
<td>Effective to increase maternal learning effects only for message which relevant to the mothers' perceived risk</td>
</tr>
<tr>
<td>George et al. (1993) India</td>
<td>Cluster trial before and after</td>
<td>Child caring practices</td>
<td>(M₁+M₄)</td>
<td>Generic messages</td>
<td>C₁ Lack of decision-making power</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C₂ Budgetary and time constraints</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C₃ Cultural inappropriateness</td>
</tr>
<tr>
<td>Campbell et al. (1994) US</td>
<td>RCT</td>
<td>Keeping pediatric preventive medical appointment</td>
<td>((M₃+M₅)+M₁)</td>
<td>Generic reminders only specified the date and time of appointment</td>
<td>C₄ Recipients had already a good understanding on perceived risks</td>
</tr>
<tr>
<td>Baker et al. (1998) US</td>
<td>RCT</td>
<td>Flu vaccination</td>
<td>((M₁+M₄)+M₂+M₃)</td>
<td>(M₁ + M₂+M₃)</td>
<td>C₆ Less precise tailoring</td>
</tr>
<tr>
<td>Jibaja-Weiss et al. (2003)</td>
<td>RCT</td>
<td>Breast and cervical cancer screening</td>
<td>(M₁+M₄)</td>
<td>M₁</td>
<td>C₇ No specific action instructions</td>
</tr>
<tr>
<td>Kreuter &amp; Stretcher (1996)</td>
<td>RCT</td>
<td>Prevention and screening behaviours</td>
<td>(M₁+M₄)+(M₂+M₅)+(M₃+M₅)+SE+(M₅+TTS)</td>
<td>M₁</td>
<td>C₈ No prioritisation</td>
</tr>
<tr>
<td>Nansel et al. (2002) US</td>
<td>RCT</td>
<td>Paediatric injury prevention</td>
<td>(M₁+M₄)+(M₂+M₅)+(M₃+M₅)+SE+Role modelling</td>
<td>(M₁ + M₂+M₃)</td>
<td>C₉ Greater adoption of the health behaviour</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C₁₀ Those who discussed the topic with their physician showed significant greater change</td>
</tr>
</tbody>
</table>

M₁ Perceived risk, M₂ Percieved barrier, M₃ Percieved benefit, M₄ personalisation, M₅ content matching, M₆ feedback, SE Self-Efficacy, TTS Trans-Theoretical Theory