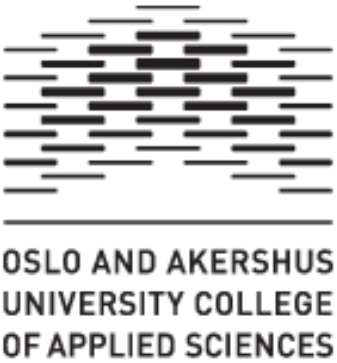
Master's thesis

Public Health Nutrition

2017

Magnus Haakens

Improving fruit and vegetable intake among Norwegian first graders through reinforcement, modelling, and peer advice: A feasibility study



Faculty of Health Sciences

Institute of Nursing and Health Promotion

Preface

I would like to thank my supervisor, Professor Kjell Sverre Pettersen, at Oslo and Akershus University College for his much needed insight throughout this process. I would also like to thank the Institute for Behavioural Analysis, especially my co-supervisor, Associate Professor Børge Strømgren, without whom this interdisciplinary project would not have been possible. Furthermore, I would like to thank Bachelor of Science (psychology) students Marie Røhme Aunemo, Thea Skau Engell, Ane Johansen Heggernes, Siri Munthe-Kass, and Katharina Risanger, who contributed with the implementation of the project. Your help was much appreciated. I would also like to thank the schools, teachers, parents, and pupils who participated in this study.

I would also like to thank Master of Science (Public Health Nutrition) student Anne Lene Løvhaug for many fruitful discussions throughout the year and Associate Professor Asgeir Brevik for always being available in his office when a question needed answering.

Finally, I would like to thank my family and significant other, Ivana Vucic, who inspired me to take on this project and contributed with bright ideas and sharp insights.

Abstract

Background: According to the national dietary survey Ungkost 3, Norwegian children's fruit and vegetable consumption is low. Several interventions have been implemented to increase consumption, but the effects have been modest. With promising effects from several western countries, the Food Dudes Programme has shown that reinforcement and role modelling can be efficient approaches to increasing children's fruit and vegetable consumption.

Objective: The main aim was to assess the feasibility of a reinforcement-, peer advice-, and modelling-based school intervention among Norwegian first graders. The secondary aim was to evaluate the preliminary outcomes of the intervention.

Method: Six first grades from four schools in the Akershus County (Norway) were recruited. The intervention used a multiple baseline design and ran over six weeks. The behavioural change components used were a token economy, teacher modelling, and peer advice. To answer the study objectives, a mixed methods approach including a questionnaire, observation, and weighted registration was used.

Results: In total, 99% of the pupils consented to participate in the study, and none of the pupils withdrew during the intervention. The teachers reported that the children understood and enjoyed the programme. Furthermore, most teachers reported that the token economy and peer advice were feasible and that the education material was satisfactory, although some improvements were suggested. The eligibility criteria, the suggested reinforcers, and the teacher modelling were found to be not feasible. Furthermore, the preliminary effects were not consistent across groups.

Conclusion: In its current state, the intervention is considered not feasible, nor is it believed to produce the desired effects. Further improvements are needed for the intervention to be considered feasible, especially regarding the choice of reinforcers and teacher modelling.

Contents

1. Background of the study 1
1.1 Delimitation 2
2. Theory 3
2.1 Health effects of fruit and vegetable intake 3
2.2 Fruit and vegetable consumption among Norwegian children 3
2.3 Norwegian fruit and vegetable interventions in schools 4
2.4 Food Dudes Programme 4
2.4.1 FDP studies. 6
2.5 PALS 7
2.6 The basis for behavioural analysis: radical behaviourism 7
2.7 Selection by consequences 8
2.7.1 Obesogenic environment 8
2.8. Consequence models of behavioural analysis 9
2.8.1 Reinforcement 9
2.8.1.1 Unconditioned reinforcers and punishers 10
2.8.1.2 Conditioned reinforcers 10
2.8.1.3 Generalized conditioned reinforcers and punishers 11
2.8.1.4 Reinforcer effectiveness 12
2.8.2 Punishment 14
2.9 Some higher-order principles of behaviour 14
2.9.1 Rules 14
2.9.2 Modelling 15
2.10 Aims and objectives of the study 16
3. Method 19
3.1 Participants 19
3.2 Inclusion and exclusion criteria 19
3.3 Materials and setting 20
3.4 Design 20
3.5 Fruit and vegetable types and weekly rotation 22
3.6 Baseline phase 22
3.7 Intervention phase 23
3.8 Changes to protocol 24
3.9 Measures 25
3.9.1 Teachers' and pupils' experiences 25
3.9.2 Researchers' experiences 26
3.9.3 Fruit and vegetable consumption 26

3.9.4 Operationalization of research questions and criteria for success 27
3.10 Data analysis 31
3.10.1 Questionnaire analysis 31
3.10.2 Qualitative analysis 31
3.10.3 Consumption analysis 31
3.11 Ethics 32
4. Results 34
4.1 Recruitment and eligibility 34
4.2 Relevance to target population 36
4.3 Intervention delivery 36
4.4 How suitable the teachers found the programme 36
4.5 How suitable the programme was for the pupils 38
4.6 Satisfaction among teachers and pupils 40
4.7 Teacher education material 40
4.8 Unexpected events 41
4.9 Adverse events 42
4.10 Outcomes 43
5. Discussion 49
5.1 Method discussion 49
5.1.1 Questionnaire 49
5.1.2 Observation and group discussion 52
5.1.3 Fruit and vegetable distribution and measurements 53
5.1.4 External validity 54
5.1.5 Design 55
5.2 Result discussion 56
5.2.1 Consent and retention 56
5.2.2 Eligibility criteria 57
5.2.3 Relevance of the intervention 57
5.2.4 Programme delivery 57
5.2.5 Suitability of the programme for the teachers 58
5.2.6 Suitability of the programme for the pupils 59
5.2.7 Satisfaction with the programme 61
5.2.8 Education material 61
5.2.9 Discussion of unexpected events 62
5.2.10 Discussion of adverse events 62
5.2.11 Outcome discussion 64
5.3 Conclusion 70

5.4 Implications and suggestions for future work 71
References 72
Appendixes 82

List of figures

Figure 1. Self-made illustration of the Food Dudes Programme behaviour change model described in Lowe et al. (2011).

Figure 2. Illustration of a multiple baseline design.

Figure 3. Overview of the phase progression for the different groups.

Figure 4. Flow of participants through the study.

Figure 5. Average weekly fruit consumption.

Figure 6. Average weekly vegetable consumption.

Figure 7. Mean difference in fruit consumption between last baseline week and maintenance phase.

Figure 8. Mean difference in vegetable consumption between last baseline week and maintenance phase.

List of tables

Table 1. Relationship between the four rules, their deliverer, and their consequences.

Table 2. Objectives and research questions.

Table 3. General objectives, research questions, operationalization, and criteria for success.

Table 4. Questionnaire items and scores measuring teachers' opinions regarding programme suitability.

Table 5. Questionnaire items and scores regarding the amounts of fruit and vegetables delivered to the pupils (answered by the teachers).

Table 6. Questionnaire items and scores on whether the time during lunchbreak was sufficient (answered by the teachers).

Table 7. Questionnaire items and scores on pupils' understanding of the programme (answered by teachers).

Table 8. Questionnaire items and scores regarding teachers' and pupils' satisfaction with the program.

Table 9. Questionnaire items and scores regarding teachers' experiences with the education material.

List of abbreviations

CI	Confidence interval
cRCT	Cluster randomized control trial
FDP	Food Dudes Programme
FVMM	Fruit and Vegetables Make the Marks
F&V	Fruit and vegetables
g	Grams
MBD	Multiple baseline design
RCT	Randomized control trial
SDB	Social desirability bias

1. Background of the study

With a high content of vitamins, minerals, and fibre, fruit and vegetables (F&V) are considered to be a part of a healthy diet, and daily intake is recommended by several national guidelines (Helsedirektoratet, 2014; U.S. Department of Health and Human Services and U.S. Department of Agriculture, 2015). Furthermore, fruit and vegetable consumption may reduce the risk of cancer, coronary heart disease, and stroke (Feldrnan, 2001; Joshipura et al., 2001; Marmot et al., 2007).

Despite the health benefits, a national survey shows that Norwegians have a low fruit and vegetable intake (Totland et al., 2012). Fruit and vegetable intake amongst Norwegian children and adolescents is also considered to be low (Hansen, Myhre, Johansen, Paulsen, & Andersen, 2016). A low intake of F&V during childhood and adolescence may be especially unfortunate, as eating habits established during childhood seems to influence future food choices (Totland et al., 2013).

To increase fruit and vegetable intake, school intervention programmes such as Fruit and Vegetables Make the Marks (FVMM) and the Pro Children Study have been implemented. While the Pro Children Study significantly increased fruit consumption compared to controls, the programme did not increase vegetable consumption significantly (Bere, Veierød, Bjelland, & Klepp, 2006; te Velde et al., 2008).

A school intervention programme with promising effects for both F&V is the Food Dudes Programme (FDP) (Horne et al., 2004). FDP is a programme that uses reinforcement and modelling to increase fruit and vegetable consumption in school children (Lowe et al., 2011). After an extended pilot study that showed positive results for both F&V consumption, the programme was rolled out nationally in Ireland in 2007 (Bord Bia, unknown). Furthermore, in 2013 Wengreen and colleagues reported a significant increase in fruit and vegetable consumption from an FDP pilot performed in the United States (Wengreen, Madden, Aguilar, Smits, & Jones, 2013).

Considering the positive results from the FDP, reinforcement and modelling appear promising in increasing fruit and vegetable consumption through school interventions. While there have been interventions conducted in Norway that use reinforcement along with social elements in other fields, to my knowledge, there has not been a nutrition intervention conducted based on the principles from the FDP in Norway (Natvig & Eng, 2012). Therefore, the aim of

the study was to assess the feasibility and short-term effects of a reinforcement- and modelling-based fruit and vegetable intervention with first graders in Norwegian schools.

1.1 Delimitation

When writing this thesis, I will attempt to follow the setup described in the master's thesis templet. However, considering that this is a feasibility study, I feel that the current master's thesis templet in itself is not satisfactory. Therefore, I will also look to the CONSORT 2010 Statement: Extension to Randomized Pilots and Feasibility Trials, because I consider this report to be the gold standard for the reporting of these types of studies.

Furthermore, due to the interdisciplinary nature of the study, the thesis is written to be understandable for individuals from the public health, nutrition, and psychological communities. For instance, the psychological section of the theory chapter is presented in a simple manner with many examples so as to be understandable to individuals with little previous knowledge of psychology.

2. Theory

In this section, I will introduce the most relevant theory and methodology for this thesis.

2.1 Health effects of fruit and vegetable intake

Regular fruit and vegetable consumption has been linked to a range of beneficial health effects. A 2007 report concluded that fruit and vegetable consumption can likely reduce the risk of cancer in the mouth, lungs, stomach, colorectum, oesophagus, and several other cancers (Marmot et al., 2007). Furthermore, the same report concluded that there was limited evidence of an inverse relationship between F&V consumption and cancer in the pancreas, ovaries, and cervix among others. In addition to reduced cancer risk, several meta-analyses of cohort studies have found an inverse relationship between fruit and vegetable consumption and cardiovascular disease (Dauchet, Amouyel, Hercberg, & Dallongeville, 2006; He, Nowson, Lucas, & MacGregor, 2007; He, Nowson, & MacGregor, 2006; Wang et al., 2014). However, a recent Cochrane meta-analysis on randomized controlled trials concluded that while the result regarding F&V consumption and cardiovascular disease was promising, more trials are needed to draw firm conclusions from experimental studies (Hartley et al., 2013). Furthermore, there are indications that fruit and vegetable intake may aid in maintaining a healthy bodyweight (Mytton, Nnoaham, Eyles, Scarborough, & Mhurchu, 2014).

2.2 Fruit and vegetable consumption among Norwegian children

According to Ungkost 3, the daily consumption of fruit, vegetables, and berries for boys and girls in the fourth grade were 193g (Grams) and 198g respectively (Hansen et al., 2016). For eighth graders, the consumption is 176g and 198g per day for boys and girls respectively (Hansen et al., 2016). Furthermore, another part of the Ungkost 3 survey for four-year-olds found an average consumption of 230g per day (Hansen, Myhre, & Andersen, 2017). For all the numbers presented above in this section, juices are limited to 100g per person per day, and potatoes are not included as a vegetable. There is no specific recommendation regarding fruit and vegetable consumption for children as there is for adults; however, there is agreement that portions should be adapted to the individual child's needs (Helsedirektoratet, 2017). For four-year-olds, the Ungkost 3 report compared their data to a 375g criterion (75% of the adult recommendation), but only nine percent consumed more than this criterion (Hansen et al., 2017). Both the Ungkost 3 report for four-year-olds and the Ungkost 3 report for fourth and eighth graders concluded that fruit and vegetable consumption is below what is believed to be ideal (Hansen et al., 2016; Hansen et al., 2017).

2.3 Norwegian fruit and vegetable interventions in schools

Several school interventions have been conducted aimed at increasing fruit and/or vegetable intake in Norway over the last 20 years (Bere et al., 2006; Bere, Veierø, & Klepp, 2005; Bjelland et al., 2015; te Velde et al., 2008). However, only elementary school interventions that uses behavioural components in addition to an either free or subscription-based programme for F&V is described in this section, as they are the most relevant to this thesis. Therefore, the Norwegian School Fruit Program and Health in Adolescence program are not described here (Bere, et al., 2005; Bjelland et al., 2015).

The intervention Fruit and Vegetables Make the Marks (FVMM) was aimed at increasing fruit and vegetable intake among Norwegian sixth graders (Bere et al., 2006). The programme was based on Social Cognitive Theory and used information regarding healthy eating through school curriculum, practical exercises in the school kitchen, and parental involvement in an attempt to create behavioural change. The programme received positive feedback from both teachers and pupils. However, the programme did not manage to increase the pupils' F&V intake (Bere et al., 2006).

Another elementary school intervention implemented in several European countries, including Norway, is the Pro Children Study (te Velde et al., 2008). The Pro Children Study used curriculum to increase knowledge and the availability of F&V in schools as well as a family component to influence F&V consumption among 10–11-year-olds. In Norway, the intervention showed promise for increasing fruit consumption, but the effects on vegetable consumption were modest (te Velde et al., 2008).

2.4 Food Dudes Programme

The Food Dudes Programme (FDP) was developed at Bangor University in Wales (Border Bia, unknown). FDP is a programme that uses reinforcement and modelling to increase the frequency of the desired behaviour. The concept behind the FDP is the three Rs—reward, role modelling, and repetition—where reinforcement and modelling help in exposing the children to F&V (Lowe et al., 2011). Then, repeated exposure helps the children get used to the taste and consistency of F&V. Furthermore, when the children are accustomed to F&V, the satisfaction of the fruit or vegetable itself maintains the increased consumption after then intervention (Lowe et al., 2011). An illustration of the FDP behaviour change model is presented in Figure 1.

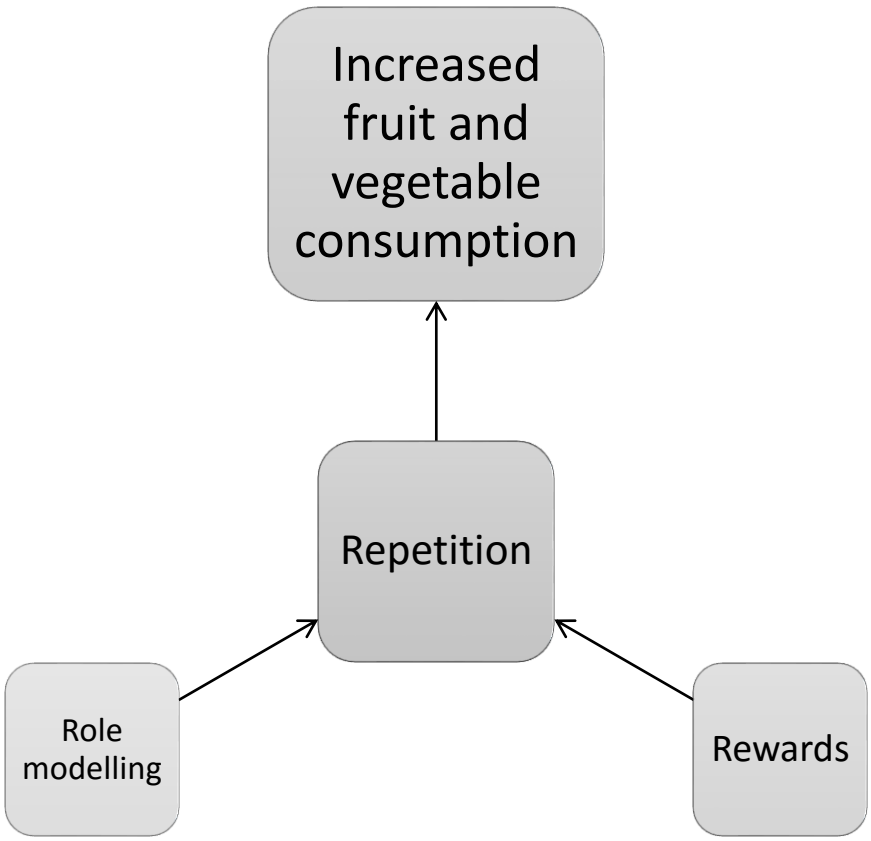


Figure 1. Self-made illustration of the Food Dudes Programme behaviour change model described in Lowe et al. (2011).

The FDP consists of two phases. Phase one is primarily implemented during school time, where the children watch videos of the heroic food dudes fighting evil to save the world. The food dude characters are children in a similar age range as the pupils in order for them to function as peer models for the children. In addition to the peer modelling, the children who eat their F&V receive a small reward and get to place a sticker on a board (Lowe et al., 2011).

Phase two of the FDP is intended to make sure that the effects of the programme are maintained. In this phase the children are encouraged to bring F&V from home in a special container, and the lowest grades continue to place stickers during this phase (Lowe et al., 2011).

2.4.1 FDP studies.

The FPD was piloted in London, and it showed promising results for both F&V consumption. Furthermore, the intervention managed to increase fruit and vegetable consumption among the children who ate very little F&V during the baseline (Horne et al., 2004). Following the successful London pilot, the programme was piloted in Ireland with similar success (Horne et al., 2009). In 2005 the Irish pilot was extended before the programme was rolled out nationally in 2007 (Border Bia, unknown).

In addition to the studies in the UK and Ireland, the FDP has been implemented with success in other western countries (Presti, Cau, Oppo, & Moderato, 2015; Wengreen et al., 2013). In 2010/2011 a pilot study of the FDP was conducted in the United States (US) (Wengreen et al., 2013). The pilot's findings were congruent with the findings from the English and Irish studies. Both fruit and vegetable intake had significantly increased when the intervention phases were compared to the baseline. Furthermore, as in the English and Irish study, the effect of the US pilot was largest among the children who consumed little F&V at baseline (Wengreen et al., 2013).

In 2012 a new study was conducted with participants who had participated in the US pilot to see how they would adapt to the new national school lunch programme. The study found that the children who had participated in the FDP prior to the launch of the new lunch programme responded better to the new programme and ate significantly more F&V compared to children who had not taken part in the FDP prior to the new national lunch programme (Wengreen, Joyner, & Madden, 2015).

A recent study also tested the FDP in Italy (Presti et al., 2015). In addition to fruit and vegetable consumption, the Italian study examined the difference between consumption of normal-weight and overweight children. The study reported that the intervention was successful in increasing home-provided F&V. Furthermore, the study found increased consumption in the intervention group when compared with the control group. The intake effect was present in both normal-weight and overweight children (Presti et al., 2015).

While there have been several successful implementations within different socioeconomic groups and countries, a study conducted in West Midlands, UK was only able to produce slight increases in fruit and vegetable consumption. Furthermore, the increase was not maintained at the 12-month follow-up (Upton, Upton, & Taylor, 2013). These findings suggest some uncertainty regarding the long-term effects of the programme. However, the majority of the studies show strong effects on both consumption and home provisioning.

2.5 PALS

Positiv atferd, støttende læringsmiljø og samhandling i skolen (PALS) is a Norwegian programme that attempts to promote positive behaviour is schools (Natvig & Eng, 2012). PALS is a multicomponent programme that uses several principles, including reinforcement and social components. Thus far, evaluations of PALS have shown promise, and the principals at PALS schools report that they are satisfied with the programme (Natvig & Eng, 2012; Sørlie, Ogden, Arnesen, Olseth, & Hansen, 2014). PALS is an example of how some of the principles used in the FDP can be helpful in promoting desired behaviours among Norwegian school children.

2.6 The basis for behavioural analysis: radical behaviourism

The philosophy of science behind behaviour analysis is radical behaviourism, based on Skinner's book "Science and Human Behavior" (1953). Radical behaviourism introduced a new scientific philosophy for studying behaviour and raised critical questions regarding the existing philosophies, such as mentalism and methodological behaviourism (Cooper, Heron, & Heward, 2014). One of the main differences between radical behaviourism and mentalism is the rejection of mental constructs as satisfying explanations of observable behaviour. For instance, Skinner would reject Freud's theory about the id, ego, and super-ego as causes of behaviour, because the id, ego, and super-ego are mental constructs.

Mentalism refers to the idea that an inner dimension exists in addition to the physical dimension and that the inner dimension can cause or mediate behaviour (Cooper et al., 2014). Skinner believed that what is referred to as a mental dimension is a part of the physical dimension. Therefore, Skinner did not draw a line at the skin; he believed that thoughts and ideas constitute behaviour, even though it is only observable to the subject itself. However, although Skinner acknowledged mental processes, he did not accept them as causal explanations for behaviour. First, the mental processes are not separate entities from the environment, and second, as there are no ways to observe mental processes, no evidence for cause and effect can be obtained. Furthermore, the goal of radical behaviourism is to predict and control behaviour, which can only be done by manipulating of the environment that controls the behaviour (Cooper et al., 2014).

2.7 Selection by consequences

The underlying idea of operant conditioning (learned behaviour) in behaviour analysis is that behaviour is selected by its consequences (Cooper et al., 2014). As behavioural analysis acknowledges the physical dimension as noted, selection by consequences means that behaviour is selected by the consequences experienced by the individual in the environment (Cooper et al., 2014). The idea that behaviour is selected based on the environment is similar to Darwin's idea about how species develop in response to their environment. It is important to note that behaviour analysis sees environment in a holistic way and not just as the immediate surroundings where the behaviour takes place. The three aspects that control behaviour—biology, learning history, and stimuli in the moment—illustrate this very well (Cooper et al., 2014). A nutritional example could be why a person just bought a hamburger. From a biological perspective, the person is prone to enjoy the taste, because high-calorie meals have been advantageous to our ancestors to survive periods when food was scarce. Due to the biological predisposition and the fact that humans often use tasty food to celebrate, the person has a learning history that links burgers to a delicious taste sensation and happiness. Furthermore, when the person walked home from work, he smelled the burgers from a nearby burger shop and saw a tasty burger advertisement on the front door. This example illustrates how it is the interaction between the individual and the environment at different levels that controls behaviour and how the aspects affect one another. Moreover, it all leads back to the overarching principle of selection by consequences.

2.7.1 Obesogenic environment

The term obesogenic environment is becoming more popular and has been studied for over 20 years (Powell, Spears, & Rebori, 2010). The term refers to an environment that promotes obesity and hinders weight loss (Powell et al., 2010). Examples of how the environment is promoting obesity is the increasingly sedentary lives we are able to live due to technological improvements, the availability of food at all times, and exposure to ultra-processed foods. Powell et al. (2010) suggested that the solution might be to focus on modifying the environment to make the environment less hostile to health. Similarly, Skinner (1953) wrote about how human biological predispositions can be disadvantageous when the environment changes. He noted how almost all humans find sugar highly reinforcing and that the biological predisposition far exceeds the biological needs in our society today. The acknowledgement that environment

is a major part of the problem, and that it is also where the solution lies, is similar to the idea of selection by consequences, at least in the sense that both fields acknowledge that behaviour is influenced by its environment.

2.8. Consequence models of behavioural analysis

While there are several principles in behavioural analysis, this thesis will be limited to two consequence models that have been rigorously researched through experiments that have accumulated a large body of evidence: punishment and reinforcement (Cooper et al., 2014). Furthermore, some principles from another class of principles called higher-order principles of behaviour, which are employed when several of the basic principles are used together, will also be described (Cooper et al., 2014).

2.8.1 Reinforcement

Reinforcement is one of the main principles in behavioural analysis. Cooper et al. (2014) use the following statement to describe when reinforcement has occurred: "If a behavior is followed closely in time by a stimulus event and as a result the future frequency of that type of behavior increases in similar conditions reinforcement has taken place" (p. 56). Two points can be emphasized regarding the previous statement. One, reinforcement always increases the frequency or probability of a given behaviour. Two, the stimulus event that results in the increased frequency occurs after the behaviour.

Reinforcement can be separated into two process categories, positive reinforcement and negative reinforcement (Cooper et al., 2014). The difference between positive and negative reinforcement is that positive reinforcement increases behaviour by introducing or intensifying a consequence, while negative reinforcement increases a behaviour by removing an aversive stimulus or reducing its intensity (Cooper et al., 2014). An example of positive reinforcement could be the social praise Bob receives from his friends for starting to eat ecologically farmed foods. Because Bob has many friends that value an environmentally friendly lifestyle, Bob receives praise and social acceptance whenever he makes environmentally friendly choices. Bob enjoys the praise and social acceptance of environmentally friendly choices, therefore he starts to purchase ecologically farmed food. Bob does indeed receive social acceptance and praise, which in turn increases the likelihood that he will continue to make environmentally friendly choices such as buying ecological food in the future, at least if his social group's interests remains the same.

To give an example of negative reinforcement we can continue to use Bob, who also bought himself a new electric car. Bob lives outside the city where he works, and therefore he has to commute to the city. Unfortunately for Bob, the US does not have the same benefits for electric cars that you find in Norway, such as no road toll. However, Bob noticed that he can avoid the toll ring if he drives off the highway and drives a mile on a side road. Because Bob is poor after spending all his money on his new car, he now experiences the road toll as an aversive stimulus. Therefore he starts to take this new commute route on his way to work to avoid the road toll. In this example Bob's behaviour of taking the side road for a mile is under negative reinforcement, as the consequence is that Bob avoids paying road tolls. As it is a common misconception that negative reinforcement results in the reduction of a behaviour, the previous statement is repeated: reinforcement always increases the frequency or probability that a behaviour will occur. Thus, it is not to be confused with punishment, which will be described later.

2.8.1.1 Unconditioned reinforcers and punishers

Stimuli that function as reinforcers or punishers without the organism having a prior learning history with them are called unconditioned reinforcers or punishers (Cooper et al., 2014). Examples of unconditioned reinforcers are food and water, and a punisher can be pain. However, even though unconditioned reinforcers do not require a previous learning history, they do not always work as reinforcers (Cooper et al., 2014). For instance, for water to have a reinforcing effect, you need have been deprived of it, or, in simpler terms, you need to be thirsty. Deprivation and other factors that influence the effectiveness of reinforcers are described later.

2.8.1.2 Conditioned reinforcers

Reinforcers or punishers that need to be learned through interaction with the environment are called conditioned reinforcers (Cooper et al., 2014). We can use the example with Bob taking the side road to avoid the road toll as an example of a conditioned negative reinforcer. Bob has learned through interaction with the environment that he only pays the road toll if he drives through the toll ring; therefore he avoids it, and thus the behaviour is under conditioned negative reinforcement.

2.8.1.3 Generalized conditioned reinforcers and punishers

Generalized conditioned reinforcers are a subgroup of conditioned reinforcers (Cooper et al., 2014). Generalized reinforcers or punishers are different from specific reinforcers such as a glass of water, because they can lead to several opportunities/reinforcers. A general reinforcer could, for instance, be money. Related to the previous example with the water glass, money can buy a bottle of water. But, money can also buy you food if you are hungry, a hotel room if you have no place to sleep, a new iPhone to make you the most popular girl in your class, or virtually anything. Because money can lead to almost anything, it is likely that it will work as a reinforcer, as it is likely that you are deprived of something it can provide, in contrast to the glass of water that will function as a reinforcer only to a person who is thirsty.

A behavioural change system that uses the advantages of generalized reinforcers is a token economy. A token economy is a system that uses symbolic reinforcers that can be exchanged for real reinforcers at a predefined ratio (Cooper et al., 2014). According to Cooper et al. (2014), a token economy system consists of three major components, as follows: (a) a specified list of target behaviours; (b) tokens or points that participants receive for emitting the target behaviours; and (c) a menu of backup reinforcers—preferred items, activities, or privileges—that participants obtain by exchanging the tokens they have earned (p. 568).

An example of the three major components in a token economy system could be an agreement between a teen and his parents regarding chores and Jet Ski availability. It could be (a) a list of chores the teen does every day when he comes home from school, such as homework, taking out the trash, and cleaning his room; (b) crosses on a board on the fridge; and (c) a list of activities such as two hours of Jet Ski time on the weekend, going to the movies, etc.

While the example is simple enough to understand, there are some aspects that are important for token economy systems to work effectively, such as exchange ratios. When establishing an exchange ratio, it is important to consider how much effort must be exerted in order to get a reinforcer. For most teen boys, riding a Jet Ski for a few hours is likely to be a potent reinforcer, but it should require a larger number of tokens for one hour of Jet Ski time than for 15 minutes of watching TV. Another point is that the exchange ratio cannot just facilitate when it is practical to deliver the backup reinforcer but must also ensure that the backup reinforcer will be delivered at a ratio that does not saturate Jet Ski behaviour.

Changing exchange ratios that change over time are also a method that is frequently used to facilitate behaviour change. When starting with the token economy system, the

exchange ratio can be low to let the individual obtain the backup reinforcer quicker (Cooper et al., 2014). Then, when the individual meets the criteria, the exchange ratio can increase. This gradual increase can be performed until satisfactory behaviour frequencies have occurred. However, it is important that the gradual exchange ratio increase is achievable for the individual.

The principles for reinforcer effectiveness described thoroughly in the following section apply to the backup reinforcer in the token economy system. Therefore, to increase the likelihood of the Jet Ski behaviour having a reinforcing effect, it should only be obtainable by the pre-specified behaviour—chores (principle of contingency). The reinforcer needs to be delivered in a satisfactory quantity, for instance one hour (principle of size). Furthermore, the individual will need to be deprived of the Jet Ski behaviour (principle of deprivation). Allowing the reinforcer to be available once a week should prevent saturation from occurring. It might seem like the token economy violates the principle of immediacy, as the backup reinforcer can be postponed. However, in a token economy system, the principle of immediacy applies to the time between the behaviour and the delivery of tokens.

In addition to the practical aspect of postponing the backup reinforcers, the increased time between the behaviour and the reinforcer also allows for more potent time-consuming reinforcers to be used. For instance, the Jet Ski backup reinforcer used in the previous example could not be delivered every day for practical reasons—there are simply not enough hours in the day. Furthermore, if the reinforcer is available on a daily basis, the teen could be saturated in regard to the reinforcer, although I question whether it is possible to saturate riding a Jet Ski for teenage boys. The ability to deliver more of the reinforcer at a lower frequency increases the reinforcer's effectiveness, because it can be delivered in larger quantities each time, which applies to the principle of size. Furthermore, the fact that time-consuming activities are likely to be less available in daily life can improve the effectiveness of the reinforcer by the principle of consistency.

2.8.1.4 Reinforcer effectiveness

Miller (2006) proposes contingency, immediacy, size, and deprivation as principles that determine the effectiveness of a reinforcer. The principle of contingency means that a reinforcer is more effective when it is delivered consistently as a consequence of a specific behaviour, and only that behaviour. If the same reinforcer is obtainable through several other behaviours, the reinforcer may not be as effective to produce the desired behaviour. For instance, if a mother

wants her child to play quietly without making so much noise, she can give her child attention while he is drawing quietly, giving him praise etc. However, if the child also gains a lot of attention by playing loudly as the mother starts to yell, he will have several ways to obtain attention. In the example above, the reinforcer is not only delivered upon the desired behaviour, thus it is likely that its effectiveness is reduced.

The principle of immediacy is defined by Miller as follows: "the more immediate the delivery of the reinforcer after the behavior, the more effective the reinforcer" (Miller, 2006, pp. 449–450). The principle of immediacy is therefore related to the time from the behaviour is performed until the consequence is presented. The longer the time between the behaviour and the reinforcer, the less effective the reinforcer will be and vice versa. If a parent praises a child for cleaning her room two days after she cleaned, the praise might not have as much effect on the behaviour as when the parent gives the praise right away. If possible, the reinforcer should be delivered within a minute, however it is not always practical to do so. The point is to deliver the reinforcer as soon as practically possible.

The next principle is size. The principle of size is related to the amount or intensity of the reinforcer delivered. An example is money on the street. If there were 50 cents lying on the street, people noticing it might just walk by, not bothering to pick it up. However, if there was a 100-dollar bill on the street, most people would not just bow down to pick it up, they would chase it around like a mad person. Certainly, size matters. There are on the other hand limits to the principle of size. For instance, if I offered a thirsty person a bottle of water for completing a simple task it is likely to have a reinforcing effect. However, if I was to offer 20 bottles, it would probably not make the person more willing to conduct the behaviour. Even though the size of the reinforcer is increased by 19 bottles, the person is well aware that one bottle is more than enough to satisfy his thirst and that water is very accessible, at least in our part of the world. The lack of effect from moving from one bottle to 20 bottles brings in the concept of deprivation, which will be described in the subsequent section.

The final factor is deprivation. The principle of deprivation means that an individual needs to be deprived of the reinforcer for such a long time that the individual would desire to obtain it. Food and drinks are good examples. If a person's favourite food is strawberry ice cream, it is likely that strawberry ice cream will function as a reinforcer for some behaviours. However, if you were to offer strawberry ice cream to the person right after a large meal, the person would be full, and the ice-cream might not function as a reinforcer at all. Another example could be that virtually anyone would perform almost any behaviour for a bottle of

water if they had gone without water for several days. But, if they were fully hydrated, they might not have any interest in the bottle of water at all.

2.8.2 Punishment

Punishment deviates from reinforcement: while reinforcement is always related to behaviour increase, punishment is always related to behaviour decrease (Cooper et al., 2014). Furthermore, punishment can also be separated into positive and negative punishment. If a mother were to yell at her children for being noisy and their noisy behaviour went decreased, the noisy behaviour would be under the control of positive punishment. The mother would add an aversive (yelling) when the noisy behaviour would occur, and the children would stop making noise. If the mother instead of yelling threatened to withdraw TV time and the children stopped their noisy behaviour, their behaviour would be under the control of negative punishment. In the last example, the mother threatens to remove a perceived good instead of adding an aversive.

2.9 Some higher-order principles of behaviour

Phenomena where several of the basic principles of behaviour work together are called higher-order principles of behaviour (Cooper et al., 2014). In the following sections I will present two principles—rules and modelling. I have limited this section to rules and modelling, as they are directly relevant to the remaining thesis.

2.9.1 Rules

According to Skinner (1969) a rule is a contingent specific stimulus. Rules can be separated into complete rules and incomplete rules. Incomplete rules are rules that do not describe the antecedent, a behaviour and its consequences, while complete rules do. There are four types of complete rules: promise, threat, advice, and warning. They can be spoken or written. What separates the rules is the following: what the consequences of the rule are and whether or not the person who delivers the rule is in control of the consequences (Skinner, 1969). A promise is a rule where the consequence is perceived to be a reinforcer, and the person who delivers the verbal instruction is in control of the consequence (Skinner, 1969). An example could be a mother telling her child that “If you put away all your toys in your room by six o’clock today, I will take you to the movies after dinner”. Advice is when the consequence is perceived to be

a reinforcer, but the person who delivers the instruction is not in charge of the consequence. An example would be if the girl's sister told her that if she put away her toys by six o'clock, their mother would take her to the movies. A threat is a rule where a person is in control of the consequence and the consequence is a perceived aversive stimulus—for example, the girl's mother saying, "If you do not put away your toys by six o'clock today, I will not take you to the movies". A warning is a rule where the consequence is perceived to be an aversive, but the person who delivers the rule is not in control of the consequence. An example would be if the girl's sister said that their mother would not take her to the movies if she did not put away her toys. Note that, in general, the more specific the rule, the easier it is to follow. Less specific rules like "If you behave well, I will buy you a treat" are often more difficult to follow, because it can be unclear what "behaving well" means. For an illustration of the four complete rules, see Table 1.

Table 1. Relationship between the four rules, their deliverer, and their consequences.

	Reinforcing consequences	Punishing consequences
-Deliverer in control of consequence	Promise	Threat
-Deliverer not in control of consequence	Advice	Warning

2.9.2 Modelling

Modelling has occurred when an individual has observed a model perform a behaviour and then performs the behaviour him/herself immediately after or at a later time (Psychology Dictionary, unknown). An example of modelling could be a mother teaching her son to tie his shoelace. The mother would first show how she ties her own slowly while the child observes. Then after a few observations, the child would manage to tie his own shoe.

According to social learning theory, observational learning is governed by four sub-processes: attention, retention, motoric reproduction, and reinforcement and motivation (Bandura, 1977). Attention means that the observer cannot learn through observation if the model does not have the observer's attention. Retention means that if the observer is going to be able to recreate the behaviour at a later point in time, he/she needs to be able to recollect the event. Reproduction refers to the need for the observer to be able to physically perform the behaviour. An example could be a child observing another child using a jump rope. The observer can observe and understand how to perform the behaviour, but she will not be able to reproduce the behaviour without training, as using a jump rope requires motoric coordination

between hands, feet, and the rest of the body. The final sub-process is reinforcement and motivation. According to Bandura, reinforcers and punishers are important, as they have a large influence on whether the observed behaviour is actually performed (Bandura, 1977). Furthermore, reinforcers and punishers can help organisms discriminate between desired and unwanted behaviour.

There have also been experiments conducted showing that some models are more effective than others. For instance, a study by Bandura, Ross, and Ross (1963a) found that children are more likely to imitate models who possess rewarding power.

I should clarify that while modelling is categorized by behavioural analysts as a higher-order principle of behaviour, the early experiments showing the effect of modelling were performed by Bandura, a social psychologist. The most famous modelling experiments Bandura performed were the Bobo Doll Experiments, which took place from 1961 to 1963 (Bandura, Ross & Ross, 1961; Bandura, Ross, & Ross, 1963b). There are differences in the scientific philosophies between behavioural analysts and social psychologists and thus how Bandura and radical behaviourists explain modelling. While Bandura acknowledges the theory of operant conditioning, he also added a mediating cognitive element between the stimuli and the response (Bandura, 1977). While a radical behaviourist would describe modelling as a discriminative stimulus, response→ reinforcer, Bandura described modelling as stimuli→ mediating cognitive element→ response. By including a mediating cognitive element as something that can be studied scientifically and can be used as explanation for behaviour, Bandura took a step away from traditional behaviouristic thought toward the mentalistic idea of cognitive psychology.

Despite the theoretical disagreements between the psychological branches on how modelling works, both agree that it influences behaviour. However, for the remainder of the thesis, I will treat modelling from a radical behaviourist point of view.

2.10 Aims and objectives of the study

In the development of the feasibility aims and objectives, the work by Orsmond and Cohn (2015) was used as a guide. While the aims and objective development was guided by the work of Orsmond and Cohn (2015), several changes were made to make the objective suitable for this specific intervention.

The main aim of this study was to assess the feasibility of the consent, retention, relevance, eligibility criteria, suitability for teachers and pupils, teacher and pupil satisfaction, teacher education material, and adverse events in a fruit and vegetable consumption

intervention among Norwegian first graders. The secondary aim was to assess the preliminary outcome trends. All objectives and research questions are presented in Table 2. To answer the study objectives, a mixed methods approach including a questionnaire, observation, and weighted registration was used.

The study did not assess the feasibility of school recruitment, intervention management, and fruit and vegetable distribution. School recruitment and intervention management were not included as objectives, because it was considered likely that the bachelor students who would be present during the intervention could bias the data needed for assessment. The distribution of F&V was not assessed, as the feasibility study used a simplified distribution system to keep costs down.

Table 2. Objectives and research questions.

General Objectives	Research questions	
Objective 1. Evaluation of recruitment capability and intervention relevance to the population	1.1 Can we recruit enough participants?	
	1.2 Is the retention satisfactory?	
	1.3 How feasible and suitable are the eligibility criteria?	A. How many children in target population are eligible for participation?
		B. Is it obvious who meets and who does not meet the eligibility requirements?
		C. Are the criteria sufficient, too inclusive, or too restrictive?
	1.4 How relevant is the intervention to the intended population?	
Objective 2. Evaluation of delivery, suitability, and satisfaction	2.1 Can the intervention be delivered as planned in the proposed setting?	A. Can the modelling by teachers be delivered as planned?
		B. Does the peer advice work as planned?
		C. Does the token economy system function as planned?
	2.2 Is the procedure suitable for the teachers?	A. Do teachers have enough time and capacity to complete the intervention?
	2.3 Is the procedure suitable for the pupils?	A. Are the amounts of fruit and vegetables presented to the pupils appropriate?
		B. Do the children have enough time to eat fruit and vegetables during the lunchbreak?
	2.4 Do the pupils understand the intervention?	
	2.5 Are the teachers satisfied with the programme?	
	2.6 Are the pupils satisfied with the programme?	
	2.7 Is the education material sufficient?	
	2.8 Are there any unexpected events?	
	2.9 Are there any adverse events?	
Objective 3. Preliminary evaluation of outcome trends	3.1 Can the study be evaluated for preliminary outcome trends?	A. Have too many adaptations been made during the intervention to evaluate outcome trends?
	3.2 Does the intervention show promise of being successful with the intended population?	A. Does examination of the data show that changes in key outcome variables have occurred?
		B. Do the changes appear to be caused by the intervention?
		C. Are the changes of the outcome variable in the expected direction?
		D. Do the effect trends suggest that the intervention has promise?

3. Method

3.1 Participants

In the recruitment process, public schools in the Northern parts of Akershus County (Norway) were contacted by e-mail regarding participation in the programme. Schools were identified through the municipality's web pages. E-mails were sent out municipality-wide, and all schools were contacted within the same week. Schools that surpassed 1.5 hours of travel time with public transportation from Oslo's central station were excluded for practical reasons. The recruitment took place from August to October of 2016.

As this is a feasibility study, a formal sample size estimation is not necessarily required. However, the sample should represent the intended population for the main study, and the sample should be large enough to inform key feasibility objectives (Thabane et al., 2010). Therefore, the sample size of the children was based on the confidence interval (CI) approach, as suggested by Thabane et al. (2010). Based on an estimated guess it was assumed that 90% of the children who received consent from their parents and consented themselves would be eligible to participate in the intervention. With a CI of 95% and the lower limit of the CI set to 0.84, it was estimated that the intervention required at least 97 children.

A large part of the feasibility objectives include teacher experience, measured by mixed methods. For qualitative research, 5–20 participants has been suggested for feasibility studies (O'Cathain et al., 2015). Considering that there are only one or two teachers per class with approximately 20 children, the sample of children would become unethically large if the study was to use a teacher sample of close to 20 teachers. Therefore, it was decided that 8–10 teachers would suffice, which would increase the sample of children somewhat but not to a large extent.

3.2 Inclusion and exclusion criteria

The inclusion criteria were the following: pupils in the first grade at one of the participating schools and parental consent to participate in the study.

The exclusion criteria were the following: limitations in cognitive function that affect social interaction, limitations in cognitive function that affect the child's ability to understand the programme, disabilities that affect the child's ability to eat by him/herself, and more than one fruit or vegetable allergy or sensitivity. Furthermore, pupils who are missing more than 30% of the total data points will be excluded from the fruit and vegetable consumption analysis.

The exclusion criteria related to cognitive function did not have a strict criterion, thus it required the teachers' and researcher's evaluation. The evaluation criteria approach was chosen in favour of strict diagnosis criteria, such as autism spectrum disorder, because the strict criteria approach could miss individuals who had not received a diagnosis yet. According to a 2012 report from the National Centre for Health Statistics, over 50% of children with autism spectrum disorder in the US received the diagnosis at five years of age or later (Pringle, Colpe, Blumberg, Avila, & Kogan, 2012). Assuming the situation is similar in Norway, a strict criterion would not be able to exclude more than about half of the individuals it was aimed at.

3.3 Materials and setting

The materials used in the study were weights, token economy boards, stickers (tokens), knives, cutting boards, a reward box, education materials, and lunchboxes. The study took place at the participating schools. A bachelor student in behavioural analysis was present at each school to help the teacher implement the intervention. The bachelor student had to be present during the lunchbreak to aid the teacher. The student was instructed to act as an assistant to imitate a normal school setting as close as possible.

In Norway, pupils do not receive lunch at school, and the common practice is to bring a lunchbox from home. Furthermore, the children eat their lunch together in the classroom. All the included schools had a standardized lunchtime of approximately 30 minutes.

3.4 Design

The study used a multiple baseline design (MBD) across school classes. Groups were set to school level and not class level, because it was assumed that classes within the same school would be more homogenous compared to other classes. The four groups were randomly assigned a number from one to four, which would determine the length of their baseline and intervention. The randomization was performed with Microsoft Excel. On the 7th of November all groups started their baseline phase. After one week of baseline, group one started the intervention. The remaining groups followed with a one-week delay between each. The study ran over six weeks and ended on the 16th of December. Figure 2 presents an illustration of an MBD.

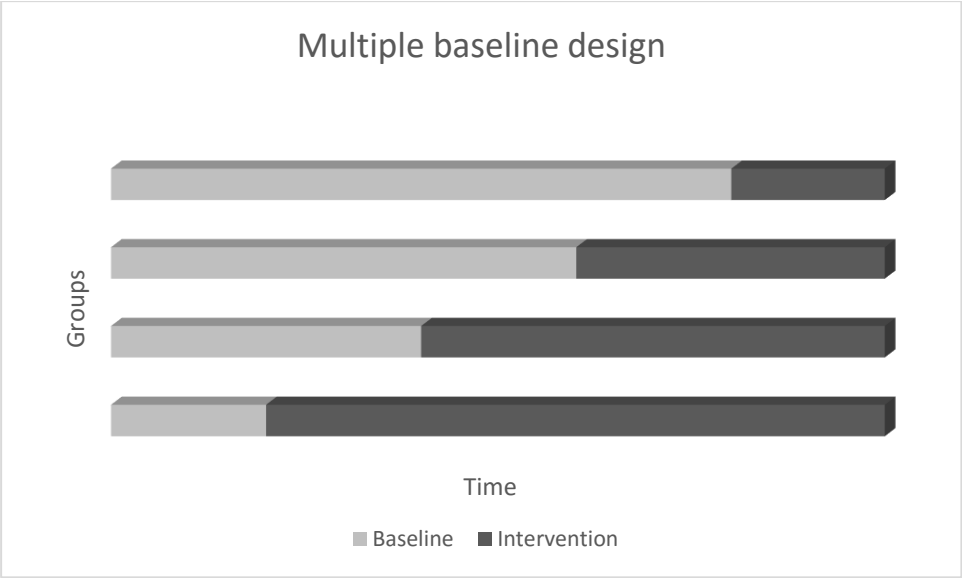


Figure 2. Illustration of a multiple baseline design.

In a MBD, two or more groups receive the same intervention at different points in time. The light grey area represents the baseline phase for all four groups. The dark grey area represents the intervention phase.

An MBD was chosen because it was considered to have some advantages over a randomized control trial (RCT) or a cluster RCT (cRCT) (Hawkins, Sanson-Fisher, Shakeshaft, D'Este, & Green, 2007). In an MBD the participants may act as their own controls, reducing the number of participants needed to obtain a sufficient sample (Hawkins et al., 2007). It should be pointed out that this was a feasibility study, hence the study was not aimed at, or powered to, assessing the effect of the intervention. However, feasibility studies may indicate trends in the data (Lancaster, 2015). Considering that the study had a small sample, an MBD would give the best effect indications, which could be useful for future work. Furthermore, an MBD also permitted all the schools to implement the programme, allowing more teachers and assistants to give feedback on their experiences. As teacher experience with the programme was deemed important for assessing programme feasibility, it was considered important to include all teachers available to increase the teacher sample.

MBD often requires more data points than a conventional RCT (Rhoda, Murray, Andridge, Pennell, & Hade, 2011). However, this was not considered a problem in this study, as we would take measures daily to gain information on how the participants responded during different phases of the intervention. Additional information on behavioural change or lack of behavioural change during the different phases was important, as it allowed an assessment of which components of the intervention worked or did not work. While a conventional RCT or

cRCT with a pre- and post-measure would be suitable for effect testing on a large scale, it would produce little information to assist further development, which is a major part of a feasibility study.

I must note that the design in this study is not intended as a test for a future study. As mentioned, MBD was chosen for the feasibility study because it allows all the groups to implement the intervention and thus be able to give feedback. However, because a larger future study would be adequately powered, a cRCT might be more prudent in that situation.

3.5 Fruit and vegetable types and weekly rotation

The fruits and vegetables included grapes, apples, pears, bananas, carrots, rutabagas, cucumbers, and red bell peppers. These eight fruits and vegetables were chosen as they are common in Norwegian cuisine, and the supplier could guaranty that they would be available during the study period. The fruits and vegetables were paired as follows: grapes with carrots, apples with rutabagas, pears with cucumbers, and bananas with red peppers. The four pairs were used in a four-day rotation in the order described above. Schools that implemented the intervention four days a week followed the same fruit and vegetable schedule as the schools who implemented the intervention five days a week. A four-day rotation was chosen to prevent the F&V pairs ending up on the same weekday each week to avoid a potential bias.

3.6 Baseline phase

A week prior to study start the teachers received the education material which described the phases and activities of the intervention. The education material is presented in Appendix 1 (original Norwegian version).

In the baseline phase, the children were presented with a lunchbox containing approximately 50g of fruit and 50g of vegetables cut into pieces at the start of their lunch. Prior to the lunchbreak the children were informed to leave any leftovers in their lunchbox and to not share food. After the lunchbreak, the students would gather their lunchboxes, and the remaining content was weighed. Teachers were also instructed not to give the children who ate the fruit or vegetables social reinforcers such as praise or attention that they normally would not receive.

3.7 Intervention phase

During the intervention phase, the children continued to receive a lunchbox of F&V at the start of their lunchbreak in addition to the following activities: The teachers and bachelor students received their own lunchboxes and were instructed to consume F&V to act as models for the children. The teachers were also instructed to encourage the children to eat F&V by giving social reinforcers such as praise or attention to the children who ate or tasted the fruit or vegetables.

In addition to social reinforcers, the children who ate the required amount of F&V also received a sticker that they placed on an individual token economy board. When a child received two tokens on their personal board, they could put a sticker on the class board. When the class as a whole had collected a pre-set amount of tokens, they received a backup reinforcer. The token boards used is presented in Appendix 2.

The schools were given a list of potential backup reinforcers that they could use. The teachers were also encouraged to use other activities that they thought could work well as end backup reinforcers. The suggested backup reinforcers were the following: reading aloud from a book, drawing/colouring, playtime in the gymnasium or other appropriate area, board games or other games the schools could provide, watching a TV show that children enjoy, singing and dancing to a video, and playtime with toys brought from home. At the start of the intervention, the children would vote for the reinforcers by a show of hands.

The first week of the intervention was an introduction phase where the required consumption of F&V to obtain an individual token was set to half a portion of fruit and half a portion of vegetables. Furthermore, the number of tokens needed on the class board to obtain the backup reinforcer was set to a moderate level where it would be possible to achieve two backup reinforcers. The criteria to obtain tokens and backup reinforcers were such that they were easily met in the first week to quickly teach the children the links between fruit and vegetable consumption, tokens, and backup reinforcers. Starting with the second intervention week, only children who ate all their F&V received a token, and the number of tokens required to obtain the end reinforcer were increased. A complete overview over the groups' planned intervention progress is presented in Figure 3.

The intervention also included a peer advice component. When a child obtained two personal tokens, he/she would qualify for a class token. The class tokens were distributed at the end of the lunchbreak. The teacher or bachelor student would call the pupils to the class board

where they placed their class token on the board while saying aloud to the class what fruit or vegetable they ate to obtain it.

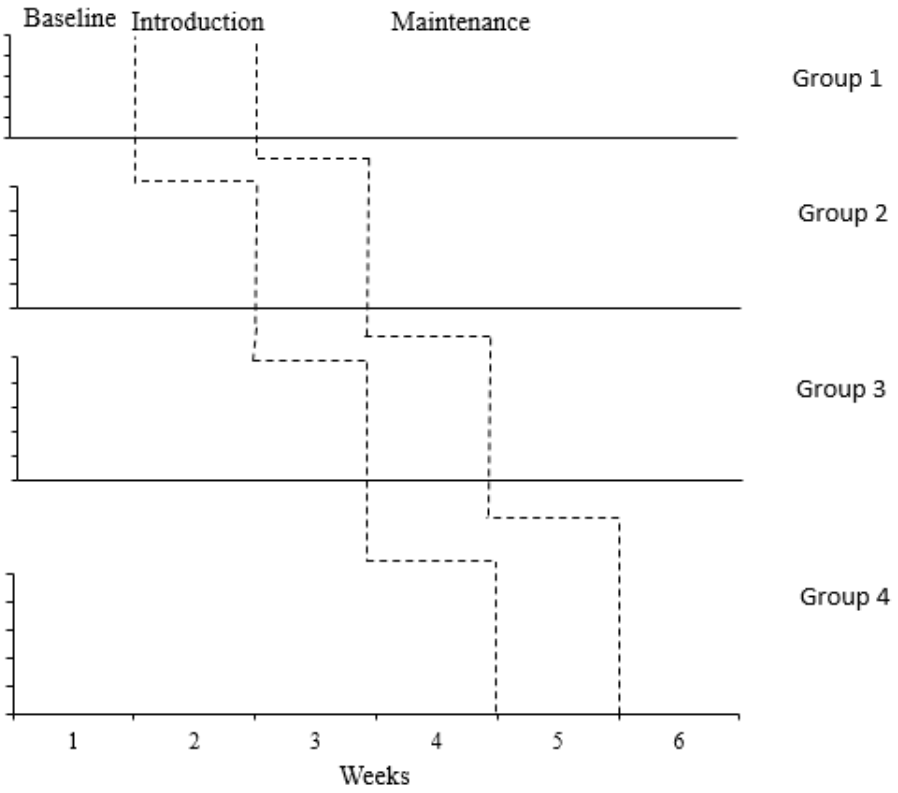


Figure 3. Overview of the phase progression for the different groups.

Figure 3 shows how the different groups were to receive the intervention. Note how the groups act as controls for eachother. For instance, at week three when Group- 1 and 2 have started the intervention, Group- 3 and 4 will still be in the baseline phase as controls.

3.8 Changes to protocol

During Group 1's first intervention week we observed that the teachers had difficulties implementing the token economy and peer advice part of the intervention. The original protocol stated that when a child had received two individual tokens he/she would receive a larger token for the class token economy board. Then he/she would go up to the class board and say aloud what he/she had eaten to obtain the token. During the first intervention week for Group 1 we observed that it was very difficult for the teacher to calm down the class for the advice activity to take place. This resulted in two outcomes: one, the child with the class token was not able to give advice due to lack of attention from the other classmates, and two, the teacher had to spend more time on calming the children down than was considered feasible for a daily activity.

Therefore, the advice part of the protocol was revised. The new protocol stated that the advice activity would take place once a week and that the pupils were to be sat down next to the token economy board during the distribution of tokens. The pupils who had earned a class token for the week would be called up to place their token and then say aloud what F&V they had eaten to obtain it. The revised protocol was implemented from week three, allowing groups 2, 3, and 4 to implement only the revised protocol.

3.9 Measures

3.9.1 Teachers' and pupils' experiences

A questionnaire was chosen as the method to gather information about the teachers' and pupils' experiences due to convenience. Interviews were considered, but the idea was discarded, because it would not be possible to conduct all interviews before the Christmas holidays, severely delaying the data collection. From an ethical aspect, interviews would also be more demanding on the teachers who had already contributed a lot of time to the study.

A new questionnaire was developed, as no appropriate questionnaire was found. The questionnaire was developed around the objectives regarding teachers' and pupils' experiences with the programme. Items were developed for all relevant objective operationalizations. For most objectives, several items were developed to reduce variability. While the items did target the teachers' attitudes toward the intervention, the items were not of a traditional psychometric nature. A Likert scale from 1–6 was used, as this is a commonly accepted attitude scale (Johnson & Morgan, 2016). An even-numbered scale was used, because it excludes a neutral option, which is hard to interpret. Even-numbered scales with 4 or 6 points also seem to have a higher reliability compared to odd-numbered scales with 5 or 7 points with a "do not know" centre point (Johnson & Morgan, 2016). Furthermore, scales with neutral/do not know options can bias the data, as individual characteristics seem to influence how often they are used (Johnson & Morgan, 2016). The questionnaire addressed the participants in the first person for the participant to answer on his or her own behalf and not teachers as a group. Several of the questions were intentionally made to be incompatible with each other to control for random checking or checking without reading the questions.

After each topic there was an optional commentary field. The commentary fields were added to include an option for the teachers to express themselves outside the specified items. As the study is a feasibility study, one aspect of interest is unexpected events that could occur

during the study. A questionnaire has a limitation in that it only asks questions the researcher already expects. The commentary field allowed the teachers to give new information, which could aid in the development of the intervention. An expert in public health and questionnaire instruments reviewed the questionnaire, and revisions were made before it was distributed. The questionnaires were distributed and collected by the head researcher the week after the intervention was complete. The questionnaire is added as Appendix 3 (Norwegian version).

3.9.2 Researchers' experiences

To measure the researchers' experience the researchers gathered within a week of the study's end to discuss their experiences. The bachelor students had been instructed to take notes on their observations during the intervention to discuss at the end of the study. The function of the group discussion was to supplement the questionnaire data and to reveal data on unexpected or possibly adverse events. The researcher experiences were intended to function as a supplement to the existing data and not a triangulation. All bachelor students had completed an introduction course in the principles of behavioural analysis and a data collection course with a focus on observation. The group discussion with the researchers took place at the Oslo and Akershus University College of Applied Sciences, and the head researcher took notes on the conversation.

3.9.3 Fruit and vegetable consumption

Fruit and vegetable consumption was measured by weighted registration conducted by the bachelor students. All the bachelor students had completed a course in data collection and had received training in weighted registration prior to the study. The weights used had an accuracy of 1g. Each lunchbox contained approximately 50g of fruit and 50g of vegetables. For practical purposes a range of 48g–52g was accepted, as it was considered too demanding for the students to cut exactly 50g.

Weighted registration was used, as it is considered one of the most accurate methods of determining food intake (Bingham et al., 1995; Perisic & Rosner, 1999). While weighted registration may be too labour-intensive to use in the intervention on a large scale, it was considered feasible for this study due to the small number of participating schools. Therefore, the fruit and vegetable measurement method used in this study will not be assessed for feasibility for a larger later study. The 48–52g range was accepted, because the data will be used on a group level with averages. In a small-scale test prior to the study, we found that the weighted lunchboxes were distributed evenly around 50g. Therefore, we expected that the

group average for lunchboxes during the study would also be distributed similarly with an average close to 50g.

3.9.4 Operationalization of research questions and criteria for success

To evaluate the intervention and research questions, three outcome criteria adapted from Thabane et al. (2010) were used: 1) feasible as is, 2) feasible with close monitoring or modifications, and 3) not feasible. For the intervention to be considered feasible, all research objectives with specific success criteria needed to satisfy criteria 1 or 2, and the qualitative evaluations would also need to be considered as satisfactory. Furthermore the effect data would also need to show promising trends. A full overview of the research questions, their operationalization, and their criteria for success can be found in Table 3. The criteria for success presented in Table 3 were used as the main evaluation tool. However if the teachers' comments or researcher group discussions revealed additional important information, this would also be taken into account.

Table 3. General objectives, research questions, operationalization, and criteria for success.

General Objectives	Research questions		Operationalization	Criteria for success
<u>Objective 1.</u> Evaluation of recruitment capability and intervention relevance to the population	1.1 Can we recruit enough participants?		Percent of pupils who received consent and consented themselves to participate in the study	-≥95% Feasible as is. -94%-90% Feasible with close monitoring or modifications. -<90% Not feasible.
	1.2 Is the retention satisfactory?		Percent of pupils who started the intervention that completed it	-≥95% Feasible as is. -94%-90% Feasible with close monitoring or modifications. -<90% Not feasible.
	1.3 How feasible and suitable are the eligibility criteria?	A. How many children in the in the target population is eligible for participation?	Percent of children in the target population eligible for participation	-≥90% Feasible as is. -89%-85% Feasible with close monitoring or modifications. -<85% Not feasible.
		B. Is it obvious who meets and who does not meet the eligibility requirements?	a. Did any pupils need to be excluded during the study? b. Did any pupils that should have been included get excluded? c. Qualitative evaluation by researchers	-a. Zero exclusions: Feasible as is. -One exclusion: Feasible with close monitoring or modifications. ->1 exclusion: Not feasible. -b. 1 Feasible with close monitoring or modifications. ->1 Not feasible. -c. No set criteria
		C. Are criteria sufficient, too inclusive or too restrictive?	Qualitative evaluation by researchers	-No set criteria
	1.4 How relevant is the intervention to the intended population?		Total mean and lower quartile mean of fruit and vegetables consumed at baseline	-Mean class consumption >30 grams of both fruit and vegetables and lower quartile mean of >10g: No need for intervention. -Mean class consumption <30 grams of fruit or vegetables or lower quartile mean of <10g: Need for intervention.

<u>Objective 2.</u> Evaluation of delivery, suitability, and satisfaction	2.1 Can the intervention be delivered as planned in the proposed setting?	A. Can the modelling by teachers be delivered as planned?	Reports from teachers or researchers of any deviations related to modelling implementation	-No deviations: Feasible as is. -Minor deviations: Feasible with close monitoring or modifications. -Major deviations: Not feasible.
		B. Does the peer advice work as planned?	Reports from teachers or researchers of any deviations related to peer advice implementation	-No deviations: Feasible as is. -Minor deviations: Feasible with modifications. -Major deviations: Not feasible.
		C. Does the token economy system function as planned?	Reports from teachers or researchers of any deviations related to token economy implementation	-No deviations: Feasible as is. -Minor deviations: Feasible with close monitoring or modifications. -Major deviations: Not feasible.
	2.2 Is the procedure suitable for the teachers?	A. Do teachers have enough time and capacity to complete the intervention?	A. Average of relevant questionnaire items summarized (negative items are reversed) into a total. Total is divided on numbers of items.	->5 Feasible as is. -5-4 Feasible with close monitoring or modifications. -<4 Not feasible.
	2.3 Is the procedure suitable for the pupils?	A. Is the amounts of fruit and vegetables presented to the pupils appropriate?	See operationalization for 2.2A	->5 Feasible as is. -5-4 Feasible with close monitoring or modifications. -<4 Not feasible.
		B. Do the pupils have enough time to eat fruit and vegetables during the lunchbreak?	See operationalization for 2.2A (teachers' answers for pupils)	->5 Feasible as is. -5-4 Feasible with close monitoring or modifications. -<4 Not feasible.
	2.4 Do the pupils understand the intervention?		See operationalization for 2.2A (teachers' answers for pupils)	->5 Feasible as is. -5-4 Feasible with close monitoring or modifications. -<4 Not feasible.
	2.5 Are teachers satisfied with the programme?		See operationalization for 2.2A	->5 Feasible as is. -5-4 Feasible with close monitoring or modifications. -<4 Not feasible.

	2.6 Are the pupils satisfied with the programme?		See operationalization for 2.2A (teachers' answers for pupils)	->5 Feasible as is. -5-4 Feasible with close monitoring or modifications. -<4 Not feasible.
	2.7 Is the educational material sufficient?		See operationalization for 2.2A	->5 Feasible as is. -5-4 Feasible with monitoring or modifications. -<4 Not feasible.
	2.8 Are there any unexpected events?		Comments or notes by teachers or students	-No set criteria.
	2.9 Are there any adverse events?		Comments or notes by teachers or students	-No set criteria.
Objective 3. Preliminary evaluation of outcome trends	3.1 Can the study be evaluated for preliminary outcome trends?	Have too many adaptations been made during the intervention to evaluate outcome trends?	Evaluation of deviations	-No set criteria.
	3.2 Does the intervention show promise of being successful with the intended population?	A. Does examination of the data show that changes in key outcome variables has occurred?	Examination of trend lines by researchers	-No set criteria.
		B. Does the change appear to be caused by the intervention?	Evaluation of consistency between groups	-No set criteria.
		C. Are the changes of the outcome variable in the expected direction?	Analysis of trend lines	-Yes: All groups show positive trends. -No: All or some groups show negative trends.
		D. Do the effect trends suggest that the intervention has promise?	Fruit and vegetable consumption measured by weighted registration	-The intervention show promise if: >10g mean difference between baseline and maintenance phase with 75% CI >10g.

3.10 Data analysis

The recruitment, eligibility, and retention data will be presented with descriptive statistics, percentages, and CI.

3.10.1 Questionnaire analysis

The questionnaire data will be presented with descriptive statistics. To analyse each category, the data within a category were merged into a single score. To reduce the data to a single score, the Likert scale of the items with negative attitude statements was inversed, and the item averages were summarized. The summarized score was divided by the number of items, resulting in a single score between one and six. In addition to the summarized scores, the mean scores for each item are presented along with the range of the respective item.

3.10.2 Qualitative analysis

For the analysis of the teachers' comments and the group discussion, a descriptive approach is taken. Because we are observing concrete behaviours and situations, a descriptive analysis is considered to be sufficient. Therefore, there will not be hermeneutic or other theoretical interpretations of the text. Furthermore, I will not account for the students', teachers', and my own preconceptions to a greater extent than what has been done so far.

3.10.3 Consumption analysis

To analyze the effect data to obtain an indication of success, the method for the evaluation of feasibility/pilot trials suggested by Lee, Whitehead, Jacques, and Julious, (2014) was used. As pilots or feasibility studies are not adequately powered, significance testing is not recommended. Therefore, the approach suggested by Lee et al. (2014) uses mean difference, several CI, and critical significance to evaluate whether the pilot shows promise for a future intervention. The CI that will be used together with the critical significance are 95, 90, 85, 80, and 75. To establish a point of critical significance, several aspects such as health improvement and cost benefit were considered. However, a report from 2005 found that, if maintained, an increase of 25g grams in 3% of the pupils would be cost efficient (Sosial – og Helsedirektoratet, 2005). Even though this is a strong argument for school interventions aimed at increasing F&V intake, I believe that an increase of 25g for 3% of the pupils is lower than what is reasonable to

expect. Thus, cost benefit was not a good parameter to use for critical significance. I then looked to medicine to establish critical significance. In medicine, it is common practice to evaluate a new treatment against the best existing treatment available (Castro, 2007). Therefore, I started to search for suitable interventions to find the intervention with best effects in a Norwegian school setting. However, my search did not find an intervention that had been conducted on the target group that measured lunchbreak consumption. The only study found was the FVMM intervention, which measured intake on a daily basis (Bere et al., 2006). Furthermore, as noted, FVMM was not successful in increasing F&V intake. With no comparison, I had to establish critical significance based on the subjective notion of what I believed was reasonable that an intervention should be able to achieve. Therefore, critical significance was set to a 10g increase for both F&V per day.

Because this is a MBD, the groups was compared with themselves. A paired t-test was conducted in Statistical Package for the Social Sciences 24 to obtain the mean difference. The t-test compared the mean of the last baseline week with the mean of the maintenance phase. These means were chosen to reduce the risk of random variation and to compensate for the fact that the intervention measured several fruit and vegetable consumption behaviours. The final baseline week was chosen over the entire baseline phase to safeguard against a possible Hawthorne effect to the extent possible. Furthermore, for the intervention to be evaluated as likely to be successful, the lower bound of the 95% CI needs to be positive, and the lower bound of the 75% CI needs to be above a mean difference of 10g. If a stricter CI than 75% is above a mean difference of 10g, the intervention will be evaluated to have a high likelihood of effect in the future. With the four groups being tested independently from each other, the criteria stated above apply to all groups, and the groups will need to show consistent results.

In addition to the mean difference and CI approach described above, weekly averages for each group is also presented in figures to show how the phases of the intervention influenced intake. In the figures the consumption data is presented with weekly averages instead of day-to-day consumption as the behaviors differ from one another.

3.11 Ethics

Informed consent was collected from both parents and teachers. The consent forms clearly stated that the parents could withdraw their child/children from the study at any time. Consent from the children's parents was collected by the teachers. The teachers distributed the consent form to the parents by the teacher-parent communication system used at their school. The

consent form given to the parents and teachers clearly stated that this was a feasibility study. In addition to the consent forms, the parents also received a questionnaire with questions regarding allergies to the F&V used in the study. At the study's start, the children were verbally informed that participation was voluntary and that they at any time could withdraw from the study if they chose to. Children who did not fit the inclusion criteria but who wanted to participate in the programme received the full programme. However, data were not collected for these individuals. Parents who did not want data collected on their children would also be given the opportunity to let their child participate in the class activities without data collection. However, no parents refrained from consenting. Including all children in the programme enabled the ethical treatment of all children and reduced the risk of data contamination. This study has been approved by the Norwegian Centre of Research Data (NSD). Pupil and teacher consent form, allergy form and NSD approval is presented in Appendix 4, 5, 6, and 7 respectively (Norwegian versions).

4. Results

4.1 Recruitment and eligibility

A total of 54 schools were contacted. Seven schools accepted the inquiry, and four were enrolled in the study. The four schools included were the first four to accept the participation request. The additional schools were excluded, because the target recruitment for both pupils and teachers had been met. The sample consisted of six elementary school classes (first grade) from four schools located in three different municipalities in Akershus County. The six classes consisted of 124 pupils and nine teachers/assistants. Of the 124 pupils, 68 were male and 56

All of the 124 children received consent from their parents to participate in the study. However, after the children were informed of their right to withdraw at the study's start, one child did. Of the 124 children approached 123 consented, resulting in 99% (CI 97.3–-100.8) of the participants consenting to participate in the study. However, the researchers' group discussion revealed that several teachers had reported that some parents from the immigrant community struggled with the consent forms due to poor Norwegian language skills.

Of the 123 children who received consent, three where excluded based on the exclusion criteria. One child was excluded based on the allergy criteria, and two were excluded based on the criteria regarding limitations in cognitive function that affected social interaction. However, the researchers and teachers did not manage to apply the eligibility criteria in the intended manner. At one of the schools, one pupil was excluded from the study after a few weeks due to tantrums during the lunchbreaks. The researchers' group discussion also revealed that the researchers and teachers had trouble applying the exclusion criteria. The criteria were too vague to be used as intended in the proposed setting. It was agreed upon that the excluded children were rightly excluded. However, there was a consensus that it was not obvious who was eligible and that the criteria needed to be simpler to use to be feasible. The researchers chose not to assess the eligibility criteria for being too inclusive or restrictive, as the new eligibility criteria needed to be developed regardless of the outcome. By including the pupil who was excluded during the study, the proportion of children who fit the eligibility criteria was 96.7% (CI 93.5–99.9).

Leaving out the pupil who had to be excluded during the study, 119 of 119 pupils who started completed the study. With no withdrawals, the proportion of pupils originally invited who completed the study was 96% (CI 92.6–99.5). The flow of the participants through the study is presented in Figure 4.

All but one teacher completed the intervention and answered the questionnaire at the study's end. The teacher who did not complete the intervention was excluded because she had sick leave during the intervention and not because she herself wanted to withdraw from the study.

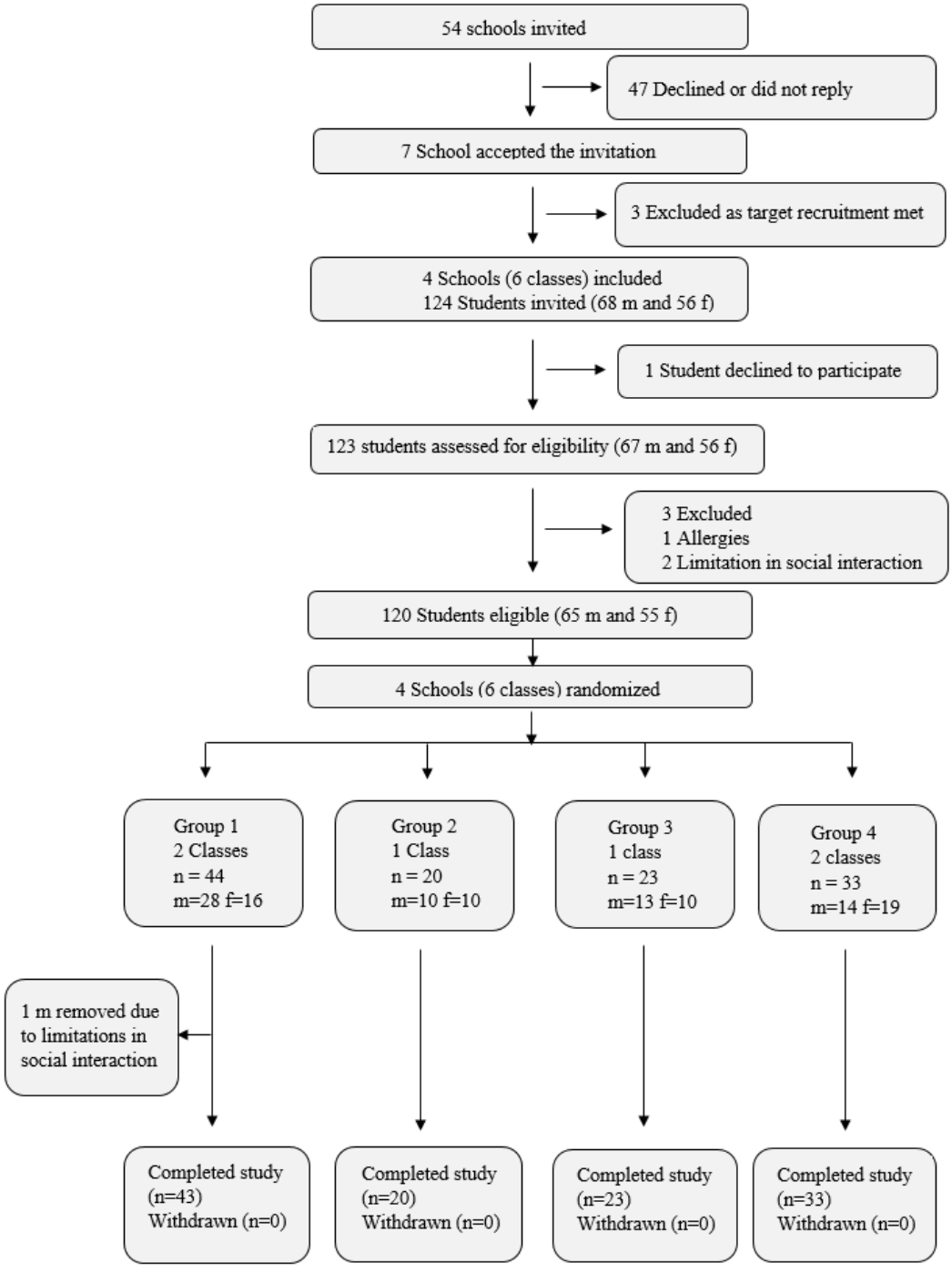


Figure 4. Flow of participants through the study.

Figure 4 shows the recruitment, exclusion, and allocation of the study. Schools with more than one first grade were allocated to the same group to reduce the risk of data contamination.

4.2 Relevance to target population

At baseline the average daily consumption of fruit was higher than the average daily consumption of vegetables. The average fruit consumptions were 34.4g, 33.1g, 23.1g, and 27.9g per day for groups 1–4 respectively. The average fruit consumption of all classes was 29.6g per day. The averages for vegetable consumptions for groups 1–4 were 22.6g, 20.3g, 16.7g, and 16,4g per day, respectively. The average vegetable consumption for all classes was 19g per day.

At baseline, the lower 25% of groups 1–4 consumed 12g, 10.8g, 11.3g, and 3.9g of fruit per day, respectively, with a combined average of 9.5g per day. Groups 1–4 lowest consumers' consumed 5.5g, 4.3g, 2.7g, and 1.1g of vegetables per day, with a combined average of 3.4g per day.

4.3 Intervention delivery

The modelling and social reinforcement performed by the teachers did not work as intended. The researchers observed that it was common practice for the participating schools to either read aloud to the children or let the children watch a learning programme during the lunchbreak. With the students watching a learning programme, the teachers did not function as models because they did not receive any attention. When the teachers were reading aloud and had the pupils' attention, they reported having a hard time eating fruit or vegetables to act as models. While the children's attention was directed toward the learning programmes, the bachelor students also observed that it was difficult for the teachers to give the children who ate or tried a new fruit or vegetable immediate social reinforcers.

After the revision of the protocol described earlier the peer advice and token economy worked as intended. The pupils who gave advice to their classes had their classmates' attention. The teachers in group one who tested the old and new protocol also expressed that the new protocol was more feasible as it fitted much better into their daily activities. However, another teacher noted a concern regarding how the token economy regarding F&V would affect other similar systems they were using aimed at reading.

4.4 How suitable the teachers found the programme

The summarized score for programme suitability by the teachers was 4.7. While the range of the positive items was small (3–6), the range of the negative items was large (1–6). The

teacher's additional comments further indicated a difference in experience among the teachers. Two teachers commented that the intervention was very easy and feasible to implement: "All the activities in the classroom were positive and feasible". However, two other teachers said that, the school had too little resources available to implement this programme. One teacher made a comment regarding the individual token boards and said, "It is too time consuming with a board for each pupil". Another teacher commented:

> Early in the first grade it would be difficult for one teacher to implement the intervention, especially if each pupil is going to have an individual token board. It would be more feasible in the second or third grade where the pupils can administrate more themselves.

The researcher discussion also revealed additional information. The students in Group 1 commented that a class teacher in their group fell ill and had to go on sick leave for most of the intervention period. The substitute teacher had more than enough to do simply keeping the class—described by the bachelor student as demanding—under control. The student also reported that the child who had to be excluded during the study due to tantrums was very demanding on the teacher and that this took away much of the time needed to run the intervention. The bachelor student in Group 2 reported a similar experience. She also had a child with special needs in her group. Some days the teacher had to sit down with him the entire lunchbreak to keep him/her from running around in the classroom. She said the following: "it was very difficult for the teacher to find time for the intervention when she had a child with special needs in the class".

Table 4. Questionnaire items and scores measuring teachers' opinions regarding programme suitability.

	n	Average score	Range
-I felt that the intervention was feasible	8	5.4	4-6
-I felt that the time required by the intervention was acceptable	8	4.8	3-6
-I think the intervention tasks was manageable	6	5.3	4-6
-I experienced the intervention as demanding, considering the school's available resources (scale reversed)	7	4.3	1-5

-I think the intervention required too much time from the school -day (scale reversed)	8	4.4	1-5
-I think that the intervention needs to be simplified to be feasible in schools (scale reversed)	8	4.3	1-6

4.5 How suitable the programme was for the pupils

Most teachers thought that the amounts of F&V were appropriate, with an average score of 4.8. None of the teachers believed that either the fruit or vegetable amounts should be increased, with average scores of 1.1 and 1.3, respectively. The scores on the items regarding reduced amounts of fruits and vegetables were also low, at 1.8 and 2.3, respectively. However, the ranges on the items aimed at reduction were quite large 1–5. One teacher commented: "The boxes with F&V were full enough". Another teacher commented: "The children normally consume more fruit than vegetables; maybe it's more important to increase the vegetable consumption?".

Table 5. Questionnaire items and scores regarding the amounts of fruit and vegetables delivered to the pupils (answered by the teachers).

	n	Average score	Range
-I think that the amounts of fruit and vegetable the pupils received was appropriate	8	4.8	3–6
-I would have increased the amount of fruit the children received if it was up to me	8	1.1	1–2
-I would have increased the amount of vegetables the children received if it was up to me	8	1.3	1–2
-I would have reduced the amount of fruit the children received if it was up to me	8	1.8	1–5
-I would have reduced the amount of vegetables the children received if it was up to me	7	2.3	1–5

The combined teacher score regarding whether pupils had enough time to consume their F&V during lunch was 3.9. On the item "I felt that the children had enough time to eat the F&V during the lunchbreak", teachers scored on the upper half of the scale, while on the item "I think that time was a limiting factor for the children's ability to finish their F&V", they scored on the lower half, accordingly. However, the teachers also scored the item "I think that an extended

lunchbreak would increase the children's fruit and vegetable consumption" on the upper half of the scale, which is contradictory to the scoring on the two other items.

In the comment section, two teachers expressed that they believed that most children had enough time with the following statements: "The majority had sufficient amounts of time" and "I felt that the pupils had the time they needed, most of the time". However, the teachers also noted that there are large individual differences between the pupils: "There is a big difference in eating speed between the pupils" and "They eat at a very different pace".

Table 6. Questionnaire items and scores on whether the time during lunchbreak was sufficient (answered by the teachers).

	n	Average score	Range
-I felt that the children had enough time to the eat fruit and vegetables during the lunchbreak	8	4.5	3–6
-I think that an extended lunchbreak would increase the children's fruit and vegetable consumption (scale reversed)	8	3.1	1–5
-I think that time was a limiting factor for the children's ability to finish their fruit and vegetables (scale reversed)	8	4.1	1–5

The teachers felt that the children understood the token economy system and that fruit and vegetable consumption resulted in a reward with average scores of 5.8 and 5.3, respectively, with a summarized score of 5.6. A teacher in Group 4 commented, "We have a similar system in the PALS programme we are running; therefore, the system is well known among our pupils".

Table 7. Questionnaire items and scores on pupils' understanding of the programme (answered by teachers).

	n	Average score	range
-I felt that the children understood the token economy system	8	5.8	4--6
-I felt that the children understood that fruit and vegetable consumption would result in a reward	8	5.3	3--6

4.6 Satisfaction among teachers and pupils

The teachers rated the satisfaction items with a mean score of 4.6. However, one of the items ranged from 2–6, as seen in Table 8. The teachers also expressed their opinion on whether the children were satisfied with the programme. The teachers scored the pupils' satisfaction with the programme as a 5 on average. The range for pupil satisfaction was 4–6 for all items. One teacher commented, "The children particularly enjoyed when they got to place a large star on the big poster".

Table 8. Questionnaire items and scores regarding teachers' and pupils' satisfaction with the program.

	n	Average score	Range
-I would participate in this programme again if given the chance	8	4.6	2–6
-I would recommend this programme to other schools	8	4.5	3–6
-I felt that the children showed interest in the programme	7	5.0	4–6
-I felt that the children enjoyed taking part in the programme	7	4.9	4–6

4.7 Teacher education material

The education material received a combined score of 4.9. Also in this case the ranges on several items were large. A teacher commented the following: "I believe that the education material was written too academically even though I managed to understand the content". Another teacher stated, "There was a lot of written material; I think this should be reduced. I also believe simple language should be used in order to be understandable to all teachers and assistants in busy everyday life".

The researchers' discussion contributed with additional information regarding substitute teachers and assistants who were not present on a daily basis. There was a consensus that even though most teachers understood the education material, it was difficult for substitute teachers who were just present for a single day to understand and implement the intervention according to protocol.

Table 9. Questionnaire items and scores regarding teachers' experiences with the education material.

	n	Average score	Range
-I experienced being unsure how to execute different intervention tasks (scale reversed)	6	5.0	1–4
-I felt I had a good understanding of the programme as a whole	7	4.9	3–6
-I experienced situations where I did not know what to do (scale reversed)	7	5.3	1–5
-I have had a clear understanding of my tasks in the programme	7	5.3	3–6
-I understood the information in the education material	7	5.0	3–6
-I think the education material was unclear (scale reversed)	7	4.7	1–4
-I think the quality of the education material was satisfactory for me to be able to implement the intervention	7	4.4	2–6

4.8 Unexpected events

During the group discussion two bachelor students reported experiences where pupils had tried to hide vegetables to obtain tokens. In one of the classes, a pupil had eaten his home-brought lunch and then tried to hide the vegetables in his lunchbox from home to make the fruit and vegetable box appear empty. In another class, a student noticed that a pupil had finished her rutabaga suspiciously fast and without any evidence had to hand out a token. Later that week a concerned parent had contacted the school wondering why her child's backpack had been full of rutabaga when she came home from school.

While the study is not aimed at addressing the fruit and vegetable distribution to schools, we did experience some issues regarding shipments worth noting. Of the six weekly shipments received, three contained errors. On two occasions, a computer error in the supplier's web-based ordering system coded kilos as crates, resulting in over 100kg additional fruit being delivered. Furthermore, in one shipment an error during packing resulting in yellow peppers instead of red peppers.

Two bachelor students also reported that some of the children received fruit or vegetables, mostly fruit, in their lunchboxes from home during the intervention. They reported that the problem seemed most prevalent at the beginning and end of the intervention. At the same time, this was not the case for all schools. The remaining students reported that they had not experienced this issue at all.

The students in groups 1, 2, and 4 reported that their schools did not adapt the backup reinforcers and that they used the backup reinforcers that were suggested. However the student in Group 3 reported that the teachers in her class had modified the backup reinforcers to suit their class's needs. The teachers had used the suggested backup reinforce "watch a TV show that kids enjoy" and added an element of choice. The teachers had written all the pupils' names on popsicle sticks, placed the popsicle sticks of the pupils who had completed an individual token board that week in a bowl, and drawn a popsicle stick from the bowl. The pupil drawn was allowed to choose what the class was going to watch. It should be noted that, even though this class did watch TV shows from time to time, it was not normal for the pupils to choose what to watch.

Furthermore, the students noted that the part where the pupils were supposed to vote on which activities were going to be used as reinforcers did not take place. This did not happen because it was forgotten, as there was so much information to give at the start of the intervention or because the teachers had not set aside the necessary time for this activity to take place.

Several bachelor students also pointed out that their schools started to watch the yearly Christmas calendar TV series during their lunchbreak starting the 1st of December. The students noted that the series was very popular among the pupils, almost to the point where they forgot to eat. One of the students said the following: "The children were so engrossed by the series that the teacher had to pause the series several time during their lunchbreak to remind the children to eat".

4.9 Adverse events

The bachelor students observed some adverse events during the course of the intervention. One student reported that a pupil in her class who was excluded due to the allergy criteria was only able to eat grapes from the selected fruits in the intervention. During the intervention, this pupil received grapes as his daily fruit. After several weeks the pupil had started to show signs of aversion toward grapes. The bachelor student said that she monitored him closely and that his consumption went back to normal after about a week, eliminating the need to apply alternative measures.

The student who had the class where a pupil had to be removed from the study due to tantrums during the lunchbreak also reported an adverse event regarding the child in question. The child, who had social problems, did not respond well when other children received token

when he did not. Seeing other children receive tokens and being told that he would not receive a token as he had not eaten his F&V acted as a trigger, inducing even more tantrums than usual.

Furthermore, several students had experienced pupils who expressed that they were torn between eating their lunch from home and the F&V they were provided. When trying to encourage pupils to eat F&V during the lunchbreak they experienced children saying, “I have to eat my lunch box so Daddy doesn’t get angry” or “Mommy says I have to finish my lunchbox”.

4.10 Outcomes

The results show that the initial fruit consumption was high, at above 30g in all groups. Groups 3 and 4, with the longest baseline, also show a negative trend before the intervention was implemented. Figure 4 shows that there was an initial increase in fruit consumption across all groups when the intervention was implemented. However, during the maintenance phase the trend lines deviate from each other, showing low consistency between the groups. While groups 1, 2, and 4 have the highest consumption during the introduction period with a reduction the following weeks, Group 3 shows the largest effects and also seems to have a stabile maintenance phase.

The trend lines for vegetables are similar to the trend lines for fruit when the groups are compared among themselves, but they also deviate in a few ways. The baseline consumption for vegetables is lower than the baseline consumption for fruit. Furthermore, most groups have a higher increase in their vegetable intake than their fruit intake. The trend in Group 1 shows little or no change during the introduction phase and shows a negative trend during the maintenance phase. Group 3 seems to have a somewhat stabile maintenance phase with an initial increase of >10g. Similarly, groups 2 and 3 also show an initial increase of >10g, but for these groups the effect drops after the initial week.

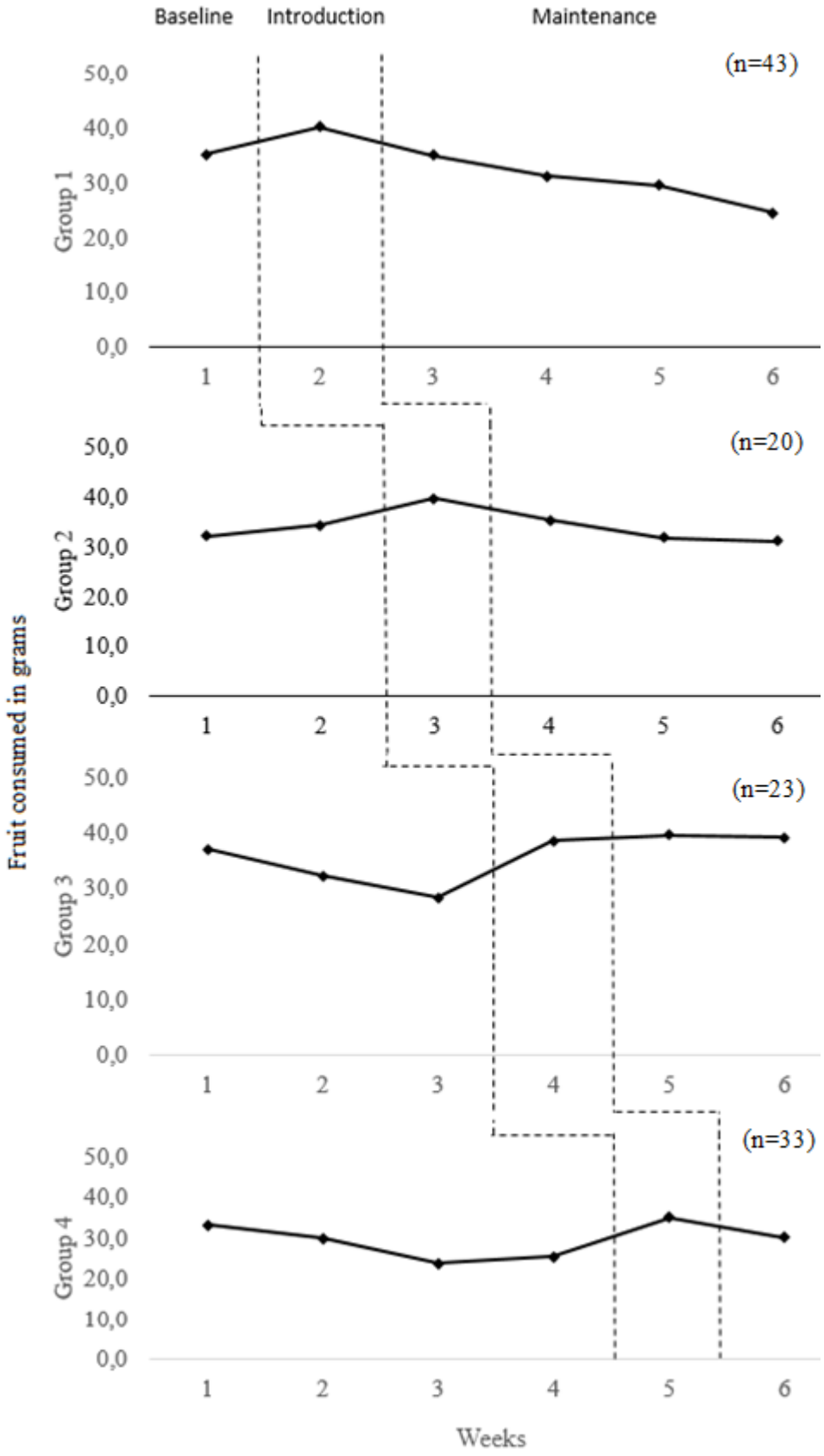


Figure 5. Average weekly fruit consumption.

Figure 5 show a slight increase in fruit during the introduction phase for all groups, but the increase is only sustained in Group 3 during the maintenance phase.

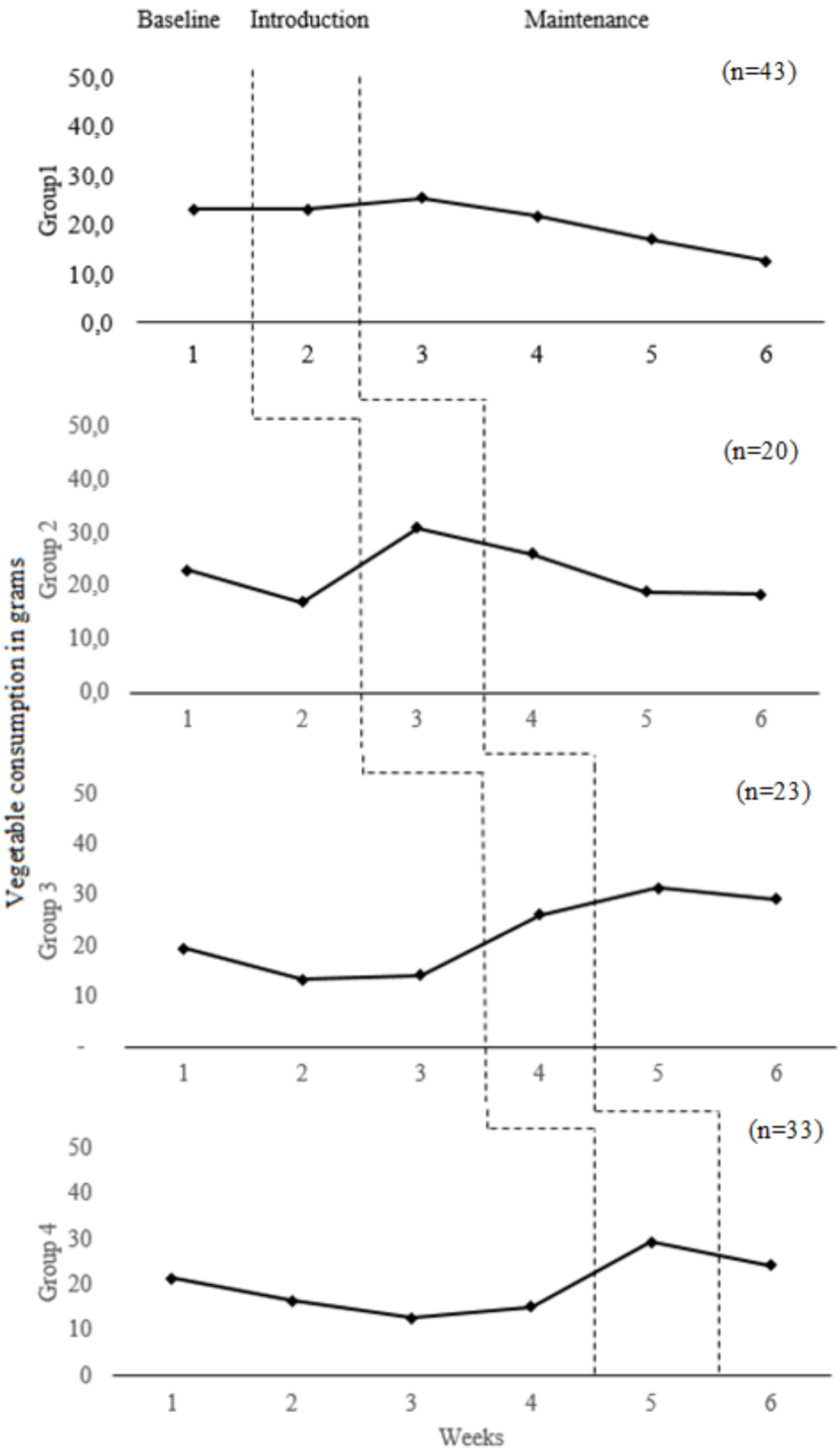


Figure 6. Average weekly vegetable consumption.

Figure 6 shows that all but Group 1 increase during the introduction phase. However, only Group 3 seems to have a stabile maintenance phase.

As there were large trend deviations between the groups, the mean differences also varied a great deal. Two of the groups, groups 1 and 2, showed negative mean differences for fruit. Group 1 had the largest negative mean difference, with the baseline week almost 5g higher than the maintenance phase, as seen in Figure 6. However, both groups 3 and 4 showed positive results. With a positive mean difference of >10g, Group 3 had the highest positive mean difference. While the lower bound of Group 3's 95% CI is over zero by a large margin, the lower bound of the 75% CI does cross the 10g mark. Furthermore, even though Group 4 shows a positive result, the lower bounds of both the 95% and the 90% confidence intervals are crossing the zero mark.

As noted, the consumption results for vegetables are slightly better than the consumption results for fruit. While Group 1 still showed a negative result, all the other groups were positive. The mean difference in vegetable consumption in groups 2, 3, and 4 was more than 5g. All the lower bounds of the confidence intervals for both groups 3 and 4 are well above zero. However, while the lower bound of the 90% CI for Group 2 is above zero, the lower bound of the 95% CI is not. Similar to the fruit data, Group 3 is the only group with a positive mean difference >10g. However the lower bound of the 75% CI is not above the 10g mean difference.

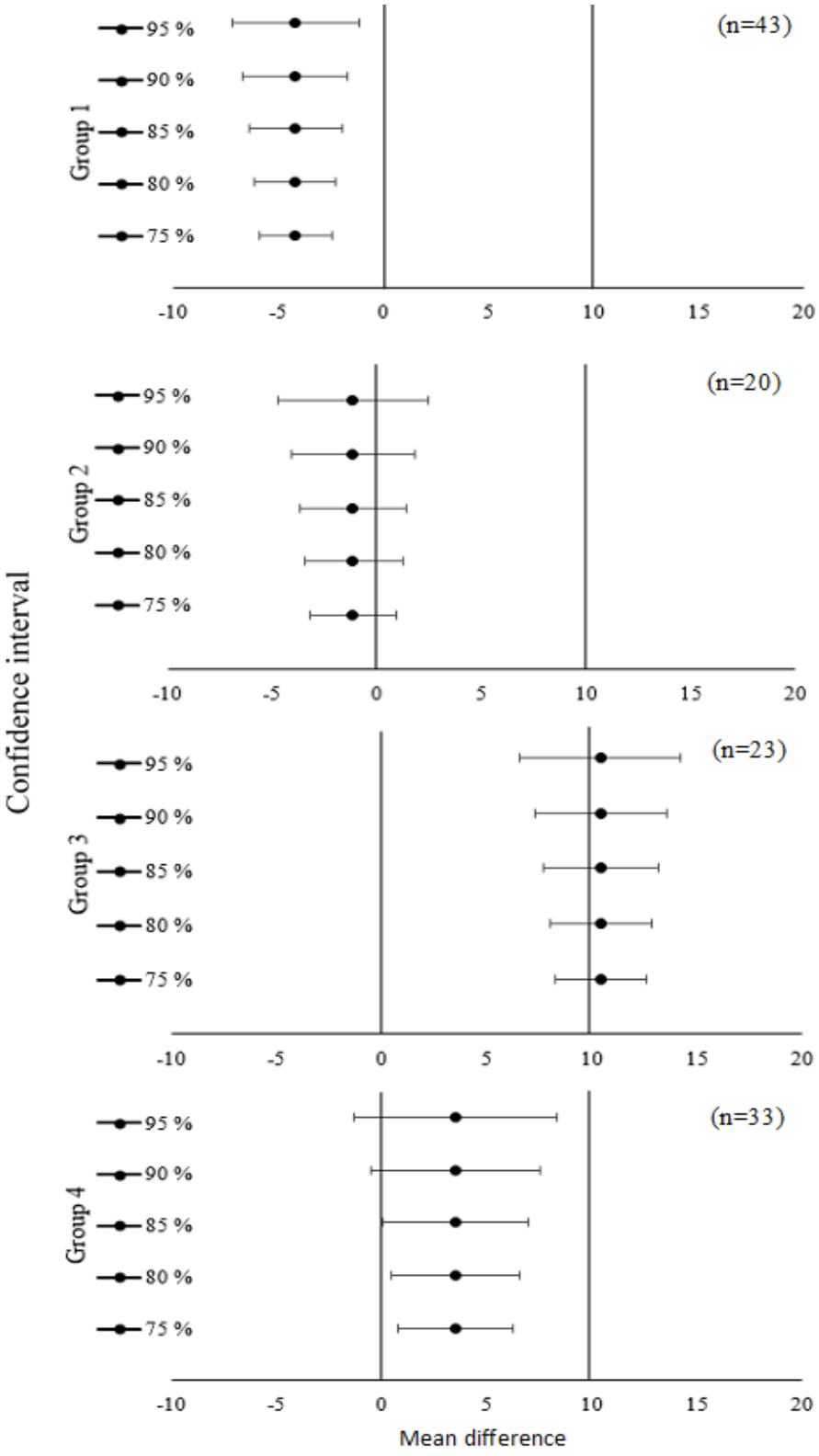


Figure 7. Mean difference in fruit consumption between last baseline week and maintenance phase.

Figure 7 show that groups 1 and 2 did not have an increase in fruit consumption from the last baseline week to the maintenance phase. Groups 3 and 4 did increase their consumption, but only Group 3 increased in fruit consumption by >10g.

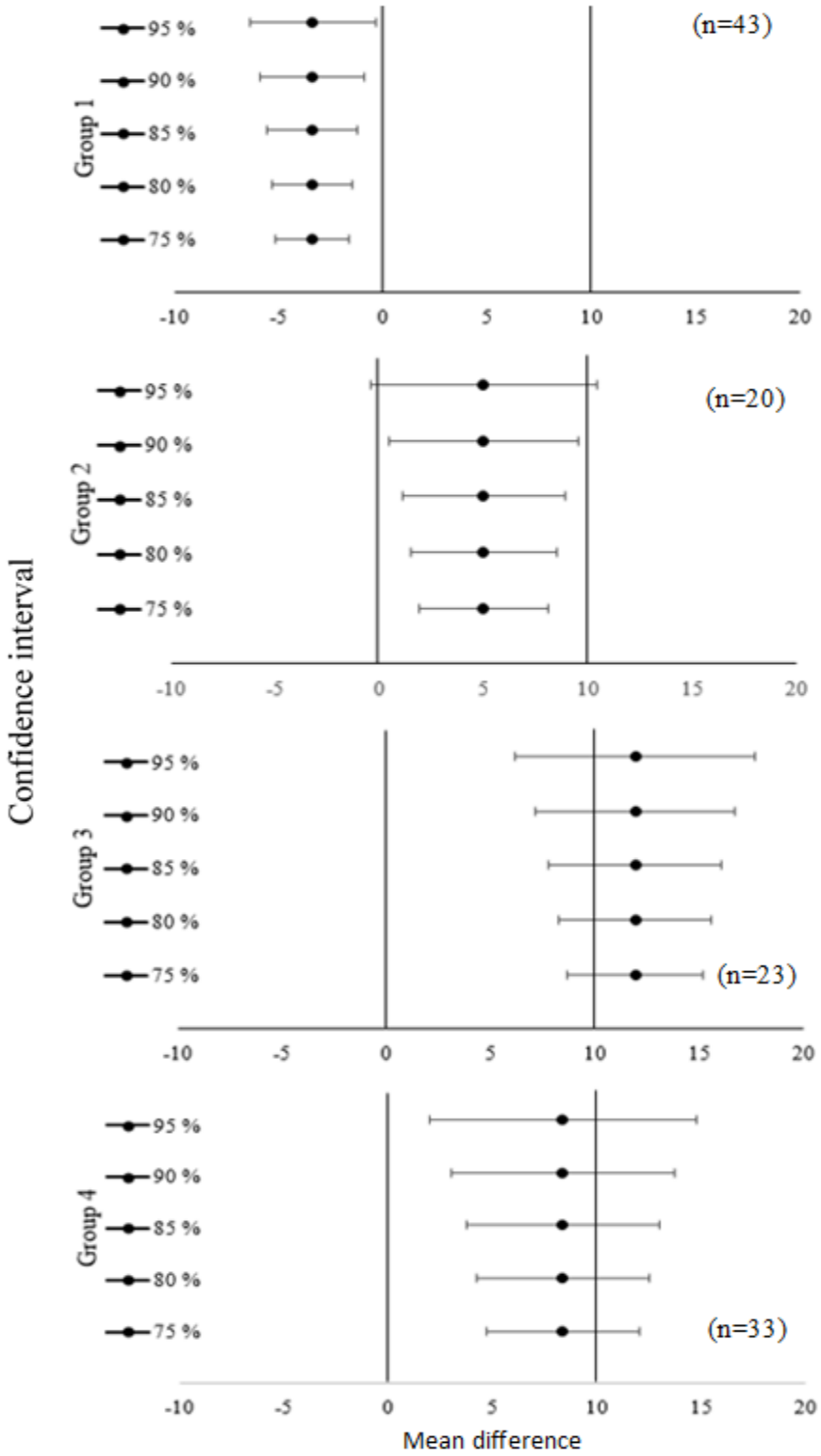


Figure 8. Mean difference in vegetable consumption between last baseline week and maintenance phase.

All groups but Group 1 increased their vegetables consumption from the last baseline week to the maintenance phase. Furthermore, Group 3 increased vegetable consumption by >10g, but the 75% CI crosses the 10g increase line.

5. Discussion

The main finding in this feasibility study was that the consent and completion proportions were high. Furthermore, most pupils and teachers seemed to be satisfied with their participation, the implementation, and the education materials although some improvements were suggested. However, the study also revealed several areas that need extensive work, such as the eligibility criteria, modelling, and suggested reinforcers. Effect results were mixed.

This study has several strengths, including a strong design, numerous data, ensured implementation, and the use of weighted registration for fruit and vegetable consumption. The study also has several weaknesses, such as different behaviours from day to day, a questionnaire that lacks validation, and participants that may not represent the target population. The subsequent section will discuss in detail the strengths and weaknesses of the methods used.

5.1 Method discussion

In this section, the data collection methods, design, and validity are discussed.

5.1.1 Questionnaire

As with all data collection methods, questionnaires are vulnerable to bias. Choi and Pak (2005) identify 48 biases in their article "A Catalog of Biases in Questionnaires". While I cannot address them all in this thesis for practical reasons, I will discuss a few. However, I do acknowledge that there are biases that are not discussed that could affect the questionnaire data. Furthermore, I will discuss the validity and reliability of the questionnaire used in the study. The discussion regarding testing will be limited to classical test theory and the threats this poses to the instrument's validity and reliability.

An important point to discuss is success criteria. With no prior data to compare to, the success criteria used for the questionnaire were based on an estimated guess. There are obvious uncertainties regarding the estimates used. Without any formal testing that could provide a reference point, the estimated success criteria used for high, medium, or low scores are uncertain. The lack of knowledge regarding success criteria further adds uncertainty regarding the questionnaire data.

In this study, an even-numbered scale without a "do not know" option was chosen, because it may be more reliable than odd-numbered scales with a neutral option (Johnson & Morgan, 2016). Furthermore a scale that does not include "do not know" or neutral options

forces the participants to answer. Forcing answers could compromise answers and thereby reduce the validity of the questionnaire. However, with the small sample available I could not run the risk of neutral or "do not know" answers that are hard to interpret.

Social desirability bias (SDB) is one possible bias. SDB occurs when people give socially desirable answers instead of their true opinions (Grimm, 2010). While the effects of SDB are most severe in a face-to-face setting, studies have also found its presence in questionnaires (Van de Mortel, 2008). Examples from health research are that people tend to report fewer calories than they actually consume (Hebert, Clemow, Pbert, Ockene, & Ockene, 1995). While SDB is difficult to eliminate, there are ways to reduce its impact, for instance by using self-administrated questionnaires rather than interview-administrated questionnaires (Kreuter, Presser, & Tourangeau, 2008). That people's answers are less influenced by SDB on self- administrated questionnaires could imply that anonymity is of importance. While the study did ensure the anonymity of the teachers and assistants who answered the questionnaires, they may not have been convinced that they were anonymous. This may be the case because the questionnaires were collected personally to save time. Theoretically, I could have named the sealed envelopes and checked the answers for each teacher. This was of course not the case; however, the teachers and assistants may have experienced it as less secure compared to mail or online surveys. If they did in fact feel that their anonymity was compromised, this could have affected their answers in an attempt to appear socially desirable.

Even though there is a risk of possible SDB at play due to the data collection procedures, I would consider this risk low. I will base this notion on the fact that the questionnaire did not collect any sensitive information. However, some of the negatively framed questions might be affected by SDB, as discussed in the results discussion.

The lack of a literacy test among the target group may have also affected the data. The questionnaire was literacy tested by my colleagues and me. However, a literacy test was not conducted among the targeted population of first grade teachers and assistants. Because the questionnaire has been tested only on educated professionals, it is possible that the literacy level required may be too high. Therefore, it is possible that the instrument could measure literacy ability rather than the intended attitudes. Furthermore, it is a possibility that groups from different parts of society understand phenomena differently and thus that the instrument measures a slightly different aspect than intended.

To try to establish content validity, an expert in health communication and I evaluated the questionnaire. While it is a strength that content validity was evaluated, it must be noted that the evaluation conducted was not ideal. The expert who was consulted is an expert in health

communications and questionnaires and could therefore evaluate the technical aspects of the scheme. However, neither the expert nor I are experts within the field of behavioural interventions or, more specifically, feasibility studies. Therefore, the expert could not conduct the evaluation of the specific item groups in relation to the topics they intended to measure. In an ideal situation, an expert panel instead should have evaluated the content validity, but given the time and resource limitations of this thesis, this was not possible.

The questionnaire was not evaluated for criterion validity. Criterion validity refers to whether an instrument performs in accordance with a reference (Bolarinwa, 2015). Alternatively, criterion validity is evaluated based on its ability to predict future performance or behaviour depending on what it is intended to measure (Trochim, 2006). Because this questionnaire was aimed at a specific intervention not previously implemented, there was no reference that it could be compared against.

Ideally questionaries should be piloted and evaluated statistically. However, due to practical considerations described later, statistical tests were not performed. Therefore, I will mention a few tests briefly and note the uncertainty the lack of these tests pose. Normally a factor analysis is undertaken, either confirmatory or exploratory, or alternatively a principle component analysis, depending on the instrument. A factor analysis is a method used discover underlying structures among the variables (Hair, Black, Babin, Anderson, & Tatham, 2006). Factor analysis is therefore a measure of construct validity. Construct validity is arguably one of the most important validities of a psychometric instrument. The lack of statistical testing with factor analysis means that the uncertainty regarding the construct validity of the instrument is high. However, this does not imply that the instrument has low construct validity, just that it is uncertain. The uncertainty means that we do not know whether the proposed items actually form the intended constructs. If the questionnaire does have low construct validity, this would be problematic, because it would mean that it does not measure what is intended.

In addition to validity there are tests developed to evaluate reliability. A common test performed to assess reliability is Cronbach's alpha. Cronbach's alpha measures the consistency of the entire scale (Hair et al., 2006). Another reliability test is test-retest. Test-retest is where participants receive the same instrument at two separate times (Hair et al., 2006). The data from both measurements are compared, and this comparison gives an indication of the instrument's stability over time. The lack of reliability tests creates uncertainty regarding the dependability of the instrument.

It is clear that there is uncertainty regarding the instrument's validity and reliability. This poses the question of why a pilot was not performed to obtain proper validity and reliability

tests. However, considering that to conduct a factor analysis one should have a variable to observation ratio of 10:1, analysis of the data was not appropriate (Hair et al., 2006). While this may be used as an argument for qualitative methods such as interviews, interviews were not possible within the available timeframe, as described in the method section. However, the questionnaire did include open commentary fields; the strength of this approach is presented in the following section.

The commentary fields allowed teachers to report additional information that the questionnaire did not address. Furthermore, the commentary fields allowed teachers to elaborate why they answered the way they did. For instance, some teachers gave the education material a medium score. Although this score gave an indication of the teacher's satisfaction with the education material, it only informs us that there is a problem, but not what the problem is. However, the teacher comments elaborated on the scores, informing us of difficult language and that the education material is too long. As shown, the commentary fields helped in answering the question of what the problem is in contrast to the questionnaire scores that gave information on whether or not there was a problem. In retrospect, given the circumstances, I consider the combination of a quantitative and a qualitative method to assess teachers' experience as suitable. However, it is important to be cautious when interpreting the results from the questionnaire due to the lmitated testing of the instrument and small sample.

5.1.2 Observation and group discussion

Validity in qualitative research refers to the "appropriateness" of the tools, processes, and data (Leung, 2015). Regarding the appropriateness of the tools, the open questions in the questionnaire and the bachelor students' notes should be suitable methods to obtain data on unexpected events. However, parts of the process do have a few shortcomings. For instance, the sampling of teachers was constrained by how many additional schools it was ethical to recruit, considering that we had obtained the required number of pupils. Therefore, it is uncertain whether or not the study reached saturation. While the study did obtain a good deal of data from the teachers, it is not unlikely that additional information could have been found if more teachers had been included in the study. Regarding the appropriateness of the analysis, I do believe that it is appropriate considering the concrete behaviours in question. However, I am sure that qualitative researchers that follow a postmodern philosophy would disagree with me on this point.

The researchers' experience was obtained from observation and daily interaction with teachers and pupils. The presence of the bachelor students at the participating schools was advantageous, as they could ensure that the programme was being implemented as intended. However, the presence of the students might also have had a negative impact on the validity of the teachers' experience. It is possible that the presence of a student at their school, aiding them, led some teachers to believe that the amount of work the intervention required appeared less than it would be withouth the students. While this uncertainty cannot be excluded I still believe that the ensured implementation and observational data improved the design overall. However, the teachers' experiences should be interpreted in the light of this uncertainty.

The findings from the observations are strengthened by the fact that the students had completed a course in observational measurements. Furthermore, the students were instructed to take notes during the intervention. Combining the notes with the fact that the intervention ran for only six weeks and that the group discussion took place the following week, the risk of recollection bias is considered low.

5.1.3 Fruit and vegetable distribution and measurements

Weighted measurement is considered one of the most accurate ways to measure individual dietary consumption and is often referred to as the gold standard (Bingham et al., 1995; Perisic & Rosner, 1999). The combination of weighted registration by trained professionals and weights with an accuracy of 1g should provide valid and reliable estimates of fruit and vegetable consumption. Furthermore, validity and reliability are further strengthened by the fact that it was trained personnel who conducted the weighing. The students who weighed the fruits and vegetables had not only had a course in data collection but they had also received training in weighted registration prior to the start of the study.

A possible threat to the accuracy of the fruit and vegetable data was that we did not deliver the pupils exactly 50g of fruit and 50g of vegetables. As mentioned, the 48g–52g range was chosen for practical purposes. This variation could potentially have skewed the data. However, as mentioned in the method section, the data were to be analysed on a group level. Therefore, based on the notion that the data would distribute themselves evenly around 50g, I do not believe that the decision to use the 48g–52g range had a notable impact on the data.

A possible threat to the fruit and vegetable data is that we did not collect data on daily consumption. Therefore, it is possible that any changes in consumption during the lunchbreak could be compensated for by reducing consumption later in the day. It is therefore unknown if

there was an actual increase overall in the children's fruit and vegetable intake. A possible way to control for the daily consumption could be to use an FFQ or 24-hour recall on all or at least some of the participants. I recommend that a future intervention collect intake data for entire days on a select group to control for this possibility.

5.1.4 External validity

Even though it is natural for feasibility studies to have small sample sizes, it is important that the sample represent the intended population for the main study (Thabane et al., 2010). Considering that the overreaching goal of most school interventions, including this one, is implementation on a national level, it is unlikely that this sample is representative. However, the study may be representative for public schools in the northern parts of Akershus County, from which the sample was drawn. It is a strength that the participation requests were sent out via mail, as the public school mailings on the municipality's homepage are frequently updated. Furthermore, the sample of invited schools consisted of most of the schools in the northern parts of Akershus. However, the exclusion of rural schools may compromise the sample.

Another potential bias is that the included schools were those that first accepted the inquiry. Considering that the participation requests were sent out over the course of a week, the municipalities who were contacted early in the week were favoured over the municipalities who received it later in the week, thus skewing the sampling toward certain municipalities. A better way to ensure the random selection of schools could have been to draw randomly from the schools who accepted or to use stratified sampling by municipalities.

While there are no official statistics on the number of research projects in public schools, it is common knowledge among researchers that there is higher pressure to participate in research projects for the schools closest to the capital. We did experience that several schools in the municipalities around Oslo declined because their first grade had already taken part in one, two, or even three other studies. Considering that this was presented as a project for a master's thesis, schools in urban areas might have turned it down for other more prestigious projects. Therefore, the sample may have been skewed toward more rural areas in Akershus and thus possibly be affected by participation bias.

While participation bias is present in most studies, this study may be particularly affected for two reasons. This study is a master's project; therefore, the project may be accepted by schools on an interest basis compared with national surveys where participation is expected. Secondly, with the study being a behavioural intervention, it demands quite a lot of time and

effort from the teachers involved. Therefore, it is a possibility that the teachers who agreed to participate did so because they have an interest in nutrition that exceeds the interest of the remaining school teacher population. However, even though there are many possible biases that could have skewed the sample, the four included schools did come from three different municipalities, giving a somewhat diverse sample.

Though it was originally planned, background characteristics such as socioeconomic status data on the pupils' parents were not collected. The lack of background characteristics eliminates the possibility to compare this study sample to other regional or national school intervention samples. Furthermore, background characteristics could also be compared with national education levels and other parameters to get an indication of the external validity of the sample.

5.1.5 Design

Multiple baseline is a rigorous design with a high amount of control, which has been suggested as an alternative to RCTs (Hawkins et al., 2007). However, in this study with varying day-to-day behaviours and a four-day fruit and vegetable rotation, certain problems arose regarding data analysis. By varying day-to-day behaviours I mean that a Monday could be apples while a Tuesday could be bananas. While both are fruit, the consumption of apples and consumption of bananas are two different behaviours, thus the two cannot be treated as the same behaviour when analyzing the data. If they were to be treated as the same, the intervention could appear highly effecting on Monday but with no effect Tuesday. However, due to Monday and Tuesday measuring different behaviours, it is likely that the difference observed between the days would be caused by the pupils' preference for apples over bananas, not the effect of the intervention. To make the data comparable, the data were compared on a week to week or on a phase basis. Weekly analysis gives better comparisons, but there is still the issue of the four-day fruit and vegetable rotation that was used in this study. As mentioned in the method section, a four-day rotation was chosen to avoid the F&V ending up on the same day each week. However, this made the F&V served one week slightly different from the next. Therefore, the analysis on a week to week basis could still be influenced by preference. In retrospect, the preference issue could have been solved by using five F&V in different orders each week.

Another issue with the reduction of data from days to weeks was that it reduced the amount of data points drastically. Few data points caused challenges regarding the establishment of a baseline. With groups 1, 2, 3, and 4 getting 1, 2, 3, and 4 baseline data points respectively, it is difficult if not impossible to know whether the groups have a stabile baseline.

Without a stabile baseline it is very difficult to get useful information out of the pre- and post-comparisons. It is possible that some baseline measures are artificially high because the pupils are excited about something new that was taking place, such as another adult in the classroom, etc. If this were to be the case, it would be likely that the intervention's effect would appear lower than it actually is. Due to the uncertainty regarding stabile baselines and inaccuracy in the measurements between weeks, within group data analysis might be slightly inaccurate. Furthermore, if some baselines are more stable than others are, between group comparisons may be compromised. If between group comparisons are compromised, the ability to draw conclusions regarding internal validity is reduced.

5.2 Result discussion

In this section, I will compare the findings with the success criteria set a priori. The findings will be presented in the same order as they were presented in the results section. Then I will include other findings such as comments or observations and discuss the findings. Finally, I will make a recommendation for a future intervention.

5.2.1 Consent and retention

While the consent proportion of 99% and retention rate of 100% satisfied the studies' criteria for success, the teachers reported that parents from the immigrant community struggled with the Norwegian consent form. While the teachers in this feasibility study managed to obtain consent from all parents due to several reminders, it is possible that less engaged teachers or teachers with high numbers of pupils from the immigrant community would not or could not find the time for such follow-up. Furthermore, if the intervention is to reduce social differences in health, it is important that a future study include people from all social classes and ethnicities.

There is also an important ethical aspect regarding consent forms. According to the declaration of Helsinki, participants must be informed of the activities they are consenting to participate in (World Medical Association, 2013). If the language in the consent form is unavailable to the receiver, it can be argued that the individual cannot give informed consent to participate in the study. To reduce the workload on the teachers, ensure the inclusion of people from all parts of the society, and ensure the ethical treatment of all participants, other possibilities such as oral consent or culture-sensitive consent forms should be explored before a future study commences.

5.2.2 Eligibility criteria

Even though the eligibility criteria satisfied some of the success criteria, they were found not feasible and would need major work to be applicable in a future intervention. By making the eligibility criteria purposely vague in order to exclude pupils who had not yet received their diagnosis, the criteria became too challenging to use in a practical manner. While it would not manage to exclude all appropriate pupils, fixed diagnosis criteria might be the rational choice as it would be more applicable.

5.2.3 Relevance of the intervention

The pupils consumed more fruit and fewer vegetables during the baseline phase than the pre-specified criteria for relevance prescribed. The findings that pupils consume more fruit than vegetables is congruent with results from other studies (Hansen et al., 2016; Hansen et al., 2017). While the results indicate that a future intervention may only need to address vegetable consumption, it might be prudent to include fruit as well. Even though the baseline fruit consumption is sufficient according to the set criteria, the data cannot say anything about what would happened to the fruit intake if a future intervention addressed only vegetables. If a future intervention addressed only vegetables, there is a possibility that an increase in vegetables could occur at the expense of fruit. Furthermore, there is uncertainty regarding the stability of the baseline data. Therefore, considering that the baseline fruit consumption was barely above the criteria, a future intervention might be better off including both F&V. Alternatively, a future pilot could explore whether a vegetable intervention would negatively affect fruit consumption.

5.2.4 Programme delivery

While the token economy and peer advice worked as intended, the modelling and social reinforcement element had some problems because the pupils were allowed to watch TV during their lunchbreak. While the TV practice makes modelling by the teachers difficult, it can be used as a digital opportunity. As seen in other interventions such as the FDP, video is a digital aid that can be used for effective peer modelling (Horne et al., 2004). Digital aids would also reduce the workload of the teachers and would fit very well into the school's daily activities. Even though there is a substantial cost involved in the development of videos, the cost may be acceptable as the videos would be easy to distribute and use on a large scale. The development of videos or other interactive digital aids should be evaluated for a future intervention.

There is also need for a discussion as to whether children should watch TV during their lunchbreak. While the effects are mixed, several studies have shown that watching a movie while one eats can affect the amount one eats (Mathur & Stevenson, 2015). Furthermore, one could argue that eating a meal together is an important part of the socialization process. However, this long and complex discussion extends beyond the scope of this thesis.

Another possibility to facilitate effective modelling could be to use popular children in the classes with large amounts of social influence. While there have been many interventions that use peer modelling, to my knowledge, there have not been interventions designed that specifically use pupils with high social status as peer models for their classmates. There are studies indicating that the use of peers with high social power might be prudent (Bandura, Ross & Ross, 1963a). It can be argued that pupils with high social power have a higher degree of rewarding power compared to other pupils. Therefore, the use of children with high social power as models could be a possibility for a future study.

5.2.5 Suitability of the programme for the teachers

The suitability of the programme for the teachers was acceptable, with a combined score of 4,7. However, the range of the answers was large (1–6), which indicates disagreement between the teachers. Why some gave suitability a low score might be explained by the information revealed in the comments and group discussion. There were comments regarding the time it took to administrate individual token boards. Furthermore, the problems with challenging, time-consuming pupils revealed in the group discussion could also contribute to explaining the low scores. Both these problems are somewhat problematic to address. If the individual boards were removed, it would distance the individual behaviour from the end reinforcer. Furthermore, it would not be possible to modify the reinforcers to incorporate an individual element as discussed in the reinforcement discussion.

The time issues regarding pupils with special needs is not something this study can solve. Pupils with special needs will be present in some classes, and they will often require more time from the teachers than your average pupil will. To solve this issue would require either more teachers present during lunch hours, or the intervention would need to be easier to implement. As the staffing at schools is the schools' responsibility, the intervention would have to adapt and attempt to reduce the workload of the intervention. It is also important to consider that schools are busy and complex environments. At times a pupil will have tantrums or something else will come up that will interrupt the intervention. Therefore, in a real-world

school setting the intervention implementation might be considered a success if it can be implemented above 70 or 80% of the time.

Another thing to note is that the positively and negatively framed items do not match each other perfectly. While a perfect match is not expected, it is peculiar that, when inverted, the negatively framed questions scored approximately one value below the positively framed questions. This discrepancy could have several explanations. It could indicate a problem with the construct. However, all the positively framed items are similarly scored, and the same goes for the negatively framed items, thus the discrepancy may be caused by something else. Another maybe more likely possibility is that the teachers find it easier to give positive feedback than negative feedback, hence a social bias may affect the data. If this were the case, the data would be skewed in a positive direction, resulting in a slightly higher score than the true value.

A future study should attempt to reduce the workload on the teachers and adapt the intervention around the school day as much as possible without compromising important behavioural change components. Furthermore, considering that this is a complex real-world setting, one might have to accept a suboptimal implementation rate.

5.2.6 Suitability of the programme for the pupils

5.2.6.1 Amounts of fruit and vegetables

The teachers scored the amount of F&V given to the pupils as acceptable, with an average score of 4.8. The teachers' scores on whether the amounts should be increased or decreased were also low, thus supporting the first statement. However, the range for the items related to the decrease were 1–5, indicating disagreement between the teachers. An argument can be made for a reduction om fruit, as suggested by one teacher. As mentioned, there is much evidence indicating that children eat more fruit than vegetables. It is possible that a reduction in fruit amounts could increase the amounts of vegetables consumed. Another modification in amounts could be a reduction in both F&V in an attempt to increase the F&V consumption among the pupils with the lowest consumption. While both alterations could work, both are based on the idea that the children received more total volume than they could consume. If we return to the teachers' score of 4.8 regarding amounts, this does not appear to be the case, at least not for the majority. As a teacher commented regarding time, there are large individual variations between the pupils. The discussion seems to come down to whether the intervention should focus on maximizing the consumption of the majority or focus on including everyone. As the underlying intent for this intervention is to improve public health and reduce differences in health between

social groups, I believe a compromise has to be made. The intervention could reduce the amounts of F&V from 50g to 40g. This reduction can be defended, as the health effects of going from 40g to 50g is debatable. It is also likely that the children who ate a lot did so because they were trained to eat F&V at home, thus they were likely eating enough on a daily basis anyway. Furthermore, there is always the possibility to increase the amounts if the pupils max out the scale.

I would also like to note that the amounts could be appropriate and that a behavioural component could explain why some pupils did not eat their F&V, thus also the solution. For instance, some of the children in this intervention would eat everything we gave them, every day, while other pupils ate little or close to nothing. This indicates that at least some of the children this age do not have a problem with the amounts delivered and that the problem may be related to their learning history. As noted, it is likely that the children who ate a lot did so because of training by their parents, and the same is likely true for the children who ate little or nothing. This means that for the intervention to work, the children would need to be gradually exposed to and incentivized to consume F&V. An effective way to increase a behaviour is to start with a low criteria and then increase gradually when the behaviour reaches the criteria (Cooper et al., 2014). This project attempted this approach to a certain degree by using an introduction week where the criteria to obtain the reinforcer were low. However, for the pupils who ate little from the start, the criteria to achieve a token may have been increased too rapidly. If they were struggling to eat half of the F&V and then the criteria increased by 100%, they could have simply fell off because the criteria became unreasonable for them. How to increase the reinforcer criteria is challenging when working with groups, because the pupils are consuming different amounts of vegetables. However, a well-designed strategy for increments of the reinforcement criteria could manage to include low-consuming pupils. A future study should attempt to improve on the introduction component used in this study, possibly using a more gradual increase over a longer period. If successful, the amount of F&V may not have to be reduced at all.

5.2.6.2 Enough time to eat?

Regarding whether or not the children had enough time to eat their F&V during the lunchbreak, the teachers' combined score was below what is acceptable at 3.9. However, this might be a result of poorly designed items rather than what their true opinions are. As noted, the answering of one item contradicts the answering of the other two. While the contradiction is hard to

explain, a possibility could be that there are large individual variations, as noted by one teacher. Depending on whether the teachers answered on behalf of the majority of the children or at an individual level, the answers would differ. However, given the lack of consistency in the findings no conclusion can be drawn. A future study has to evaluate the aspect of time, keeping in mind that there might be large individual differences in the target group.

5.2.6.3 Did the children understand the programme?

The teachers' combined score regarding the children's understanding of the programme (token economy) was excellent, with a combined score of 5.5. This finding is supported by other studies showing that five- to six-year-old children understand simple token economy systems (Filcheck & McNeil, 2004). With no indication that the intervention is too complex for the pupils, a future study can use a token economy system with both an individual and a class board.

5.2.7 Satisfaction with the programme

The combined scores for teacher and pupil satisfaction were 4.6 and 5, respectively, the ratings thus being acceptable and good. The wide range in the scoring could be explained by the issues some of the teachers had with the implementation. It is not unlikely that the teachers who struggled with the workload, time constraints, children with special needs, or a combination did not have a positive experience and therefore did not wish to participate again. High teacher satisfaction might be important if a future study is to be implemented according to protocol.

5.2.8 Education material

Most teachers seemed fairly satisfied with the education material, and the combined score was 4.9. However, three improvements became apparent for easier use: 1) shorten the length of the education material, preferably by using bullet points to make the information lucid, 2) adapt the language in the education material to the intended population, and 3) create a short version for substitute teachers and assistants who are only present in the class for a few days. A future study should attempt to meet these suggestions and preferably conduct a literacy test of the education material with the target population.

5.2.9 Discussion of unexpected events

An unexpected event was that some pupils tried to hide their vegetables to obtain tokens. Even though these things are hard to control, informing future teachers about these incidents might help them to stay alert. While I believe that the problem is small, it is important to deal with it, because it could create ethical problems due to the unfortunate reinforcement of inappropriate behaviour.

As it was not one of the objectives for this study, the errors in distribution will not be discussed. However, it might be prudent for a future study to consider the importance of accuracy, punctuality, and stability in addition to price when choosing a supplier.

Another unexpected event was the F&V sent from home. While it is hard to avoid this problem completely, additional information may help reduce it. Therefore, a future study should emphasize to the teachers that it is very important that the parents remembers the study start and finish dates, maybe by issuing several reminders.

Another factor that could have influenced the data is underreporting. While weighted registration performed by trained professionals should keep errors such as under or over reporting low, we did experience problems in the school environment that could cause underreporting. As noted, some parents packed fruit or vegetables, mostly fruit, as a part of their child's lunchbox. With some children consuming additional fruit that was not weighted, the actual intake could be higher than what was reported. However, with little consistency through the intervention or between groups, it is hard to interpret the effects of the F&V sent from home. While the amounts and effects are unclear, it is possible that a slight underreporting of fruit during the start and end of the intervention did occur.

The Christmas calendar TV show that engrossed the children during the lunchbreak could also have affected the data. While the data could be affected by the popular TV show, the Christmas season comes along every year. If the intervention is going to be effective in all seasons, the intervention should plan for seasonal variation if necessary.

5.2.10 Discussion of adverse events

The child who showed signs of aversion toward grapes during the intervention also revealed a gap in the protocol. While he was included in the programme to not feel socially excluded from the class, the protocol did not manage to address the complexity regarding diversity in fruit consumption. If he were to develop an aversion, it would be particularly severe as it would eliminate one of the few fruits he could eat from his diet. Therefore, the protocol should be modified to take fruit and vegetable diversity for individuals with allergies into account. A

possible solution could be to increase the variety of F&V used in the intervention from 8 to 12. Alternatively, pupils who are excluded due to the allergy criteria would skip the fruit or vegetables on the appropriate days.

Another adverse event was the pupil who had to be removed from the intervention due to tantrums when others received tokens and he did not. While it was not uncommon for this child to have tantrums in other situations, it was considered likely that the intervention created an environment that further induced tantrums. With the knowledge that the intervention could act as a trigger for pupils with social problems, it would be unethical to move forward without making sure that these pupils would be attended to in an appropriate manner. Therefore, an assessment of the school's ability to care for children with special needs in these specific situations should be undertaken. If found lacking, teachers who will participate in a later study should receive training or at least information on how to act to ensure the ethical treatment of children with social problems.

As mentioned, some of the pupils expressed concerns regarding not being able to finish their home-brought meal, as they had been instructed from home to finish their meal. This may indicate that the children have competing interests regarding what to consume during school lunch. It is likely that the children at some point have received positive or negative feedback from their parents on their lunch consumption. Thus, the consumption of their home-brought lunch may be positively or negatively reinforced from home. For the pupils with parents who do not reduce the size of their lunchbox in order for them to eat both, this may lead to a dilemma between the F&V and their home brought meal. The children might want to eat the F&V to achieve tokens but at the same time finish their lunchbox to receive praise or to avoid negative feedback from home. It might be obvious for adults who are participants of a study to alter the social environment in such a way that these social norms no longer apply. However, this may not be obvious to small children, who have not yet learned how the social context is affected by changes in the environment.

There are ways we could alter the information about the intervention to keep the children from experiencing it as a dilemma between home and school. The parents could receive additional information aimed at reducing the size of the lunch during the intervention period. Another way could be to supply parents with information regarding how their child might react and inform them that they should sit down and have a talk with their children. Regardless of what approach is taken, a future intervention should incorporate mechanisms to avoid the unnecessary stress this dilemma may induce for the pupils.

5.2.11 Outcome discussion

In this section, I will discuss whether the findings suggest that the intervention shows promising effects. First, I will answer the research questions related to objective 3. Then I will discuss relevant theory that could contribute to explaining the findings.

In preliminary studies it is not uncommon for adaptations to occur during the intervention, thus compromising the data in such a way that evaluations are not feasible. In this study the only adaptation made was the altering of the peer advice in Group 1. Because this adaptation is only a slight alteration regarding the timing of the peer advice and that it only affected the first intervention week for Group 1, I believe that all groups can be included in the examination.

As seen in figures 7 and 8, the data indicate that changes in fruit and vegetable consumption did occur during the intervention period. However, the graphs do not show the same pattern across the intervention. While all groups increased fruit and vegetable consumption during the introduction phase, there were clear deviations in how the groups behaved during the maintenance phase. The agreement between the groups in the introduction phase suggests that the intervention had a small immediate effect. However, the deviations between the groups in the maintenance phase make the data hard to interpret. There were certainly no clear indications of what the effects of the intervention would be after the first week. Furthermore, none of the groups met the success criteria of a positive mean difference with the lower bound of the 75% CI >10g. With the effect trends showing only moderate effects in selected groups and a large group variation, the findings in this study do not indicate promising effects for either fruit or vegetable consumption for a future intervention.

Even though the effects were not consistent, the intervention seemed to have had a better effect on vegetable consumption compared to fruit consumption for all groups. The improved effect on vegetable consumption could be explained by the fact that the baseline fruit consumption was higher than the baseline vegetable consumption. However, the findings that the intervention is more effective in promoting vegetable than fruit consumption contradict the conclusion of a 2012 meta-analysis which found that “School-based interventions moderately improve fruit intake but have minimal impact on vegetable intake” (Evans, Christian, Cleghorn, Greenwood, & Cade, 2012, p. 889).

The lack of results in some groups and different trends between the groups could be explained by the lack of a stable baseline, as discussed previously. Furthermore, the slight difference in measured behaviours noted previously could also be a part of the explanation. If we look at the day-to-day vegetable consumption in Appendix 8 we can for instance see that

the children seem to prefer vegetables other than rutabaga. This preference difference would suggest that weeks two and six of the intervention would yield slightly lower effects due to two rutabaga days compared to one the other weeks. An example of preference influencing the data could be the lack of initial effects on vegetable consumption in Group 1. We can see that all the other groups increase their intake during the introduction phase except Group 1. Furthermore, Group 1 introduced the intervention in week two, which had two rutabaga days. Therefore, if instead of looking at the weekly averages, we can compare the individual vegetables with their past consumption during the previous week. If we do that, we see that the vegetable consumption for all vegetables increased by between 0.8g and 9.6g. With an increase in all vegetables, it seems likely that it is the two rutabaga days keeping the weekly average from increasing. This example shows that the four-day rotation of F&V does allow preference to influence the weekly average data. However, considering that rutabaga seems to deviate more from the average than other fruits or vegetables, the influence of preference on the remaining weeks should be smaller.

Another factor that could have impacted the data is possible outliers. With the large variation, it became very difficult to establish whether a data point that deviated from the remaining points was an actual outlier or just a part of the random variation. For instance, if we look at the day-to-day fruit intake data from Group 3 (Appendix 9), we see that the last data point before the intervention began seems to deviate from the remaining points. But, because there is a large variation and that Group 3 did not have a data point on the previous day, it is hard to know whether or not it is an outlier. If it had been removed, Group 3's final baseline week would have looked similar to its second baseline week. Furthermore, the mean difference between the baseline and intervention would have been somewhat reduced.

An aspect that could explain the decline in fruit and vegetable consumption behaviour in groups 1 and 2 is season. As the intervention progressed, the schools started to include increasingly more Christmas activities. Some of the activities were food related, such as the baking of gingerbread cookies. Furthermore, several pupils were observed to have Christmas snacks in their lunchboxes such as a small chocolates or clementines. I consider it to be likely that if a pupil has the choice between the immediate sensation of chocolate and a possible activity at the end of the week if she eats her rutabaga, she would choose the chocolate every time. Therefore, it is possible that the suggested reinforcers were outcompeted by other stimuli in the environment. The increase in activities and snacks as we closed in on Christmas could contribute to explaining why groups 1 and 2 continued to drop. However, groups 3 and 4 did not show the same trend. This creates two options. One, the season did not affect the

consumption in a major way. Two, season does have an effect on consumption, but groups 3 and 4 were under a different stimuli control than groups 1 and 2. It is hard to establish which is more likely considering that we do not have data on the numbers of Christmas activities and treats for the different schools.

Another possible explanation regarding the varying trends between the groups could be that the groups are simply different from each other. For instance, this study did not control for socioeconomical status between the groups. Furthermore, Group 4 deviated from the other groups because it is a part of the PALS programme. It is possible that the previous exposure to similar programmes for teachers and pupils could have increased the effect, thus explaining why Group 4 responded better than some other groups.

5.2.11.1 Backup reinforcers

Theoretically, the backup reinforcers were one of the main elements intended at maintaining the behaviour level high at the maintenance phase. Therefore, the lack of consistent results across groups in the maintenance phase may be a result of the suggested backup reinforcers not working as intended. This may have occurred for several reasons. As noted from the bachelor students, the pupils had a lot of time during the week to play, often using the same activities as those suggested as potential reinforcers. If the pupils were exposed to, for instance, drawing on a regular basis, it is not unlikely that the pupils' drawing behaviour could be saturated, thus the backup reinforcer would have reduced effectiveness due to a violation of the principle of deprivation. Furthermore, because the pupils had access to several of the suggested reinforcers by other means than fruit and vegetable consumption, there is also a problem related to the principle of contingency. In retrospect, the suggested reinforcers should have been tested more thoroughly. However, within the given timeframe this was not possible.

An alternative to suggesting reinforcers could be to give the teachers along with the pupils the responsibility to find reinforcers for their class. According to Strømgren and Sørheim (2015), teachers in classes with normally functioning pupils often have a good grasp on what they may like as a reward. However, as pointed out by Grattan and Demchak (2014), to increase the likelihood of finding good reinforcers the children should be included in the decision. Therefore, a combination of teachers and pupils finding the reinforcers might be prudent. While the current intervention did encourage the teachers to use reinforcers they thought would work well in their class, our experience was that when provided with a suggestion most teachers would use the default option. Furthermore, the activity where the children were going to choose

their own reinforcers did not take place. If a future intervention were to remove the suggested reinforcers, the feasibility of teachers and pupils choosing their own rewards, and the resulting reinforcers, should be evaluated. Including the children in the selection process should not take more than a school hour where they suggest and vote for rewards. Even if this study failed to conduct this activity, I believe that it is possible with better planning. However, I do acknowledge that teachers might oppose this change, as it would increase the time consumed by the programme.

While most teachers used the suggested reinforcers during the intervention, we did observe a positive effect from teachers tailoring the reinforcers to their class. As mentioned, the teachers of Group 3 did modify one of the suggested reinforcers. Group 3 also seemed to respond better to the intervention than most groups. This could be a coincidence, but the tailored reinforcer used in that particular class could also contribute to explaining the improved effects. The use of popsicles and choice added two new elements to the suggested reinforcers. One, the use of the popsicles increased the value of the reinforcer, because in order for the pupils to have a possibility to choose what the class would watch, they would need to complete at least one individual board that week. This created an individual element. A problem with the class-based activities was that some children could chose to not eat fruit or vegetables and still obtain the end reinforcer. These children would not really have an incentive to eat, as they obtained the activity regardless of their consumption. While the popsicle method used in Group 3 did not exclude them from the end reinforcers, it did exclude them from the possibility to choose what to watch. The second element was that one of the pupils had to choose what to watch. As mentioned, it was not common for the pupils to choose what to watch. This created an element of choice, something the pupils did not have before. This could have increased the value of the reinforcer, because now the pupils who ate would not only get to watch a show but have an opportunity to watch what they wanted to watch. Furthermore, by modifying the activity, it is possible that the daily or weekly activity of watching a series which would not work as a reinforcer due to saturation was turned into something new and interesting that was only obtainable in relation to the intervention that did work as a reinforcer. Whether a future study suggests reinforcers or relies on teachers and pupils to develop them, the inclusion of an individual aspect should be attempted.

Another discussion regarding reinforcers is the type of reinforcers a future study should use. Both activities and tangible rewards can function as reinforcers. If it is the case that the lowest grades have a lot of free playtime, it could be challenging to find activities that will function as reinforcers. Alternatively, as shown in the FDP, tangible reinforcers can also work

well. Tangible reinforcers have some advantages over activities. The pupils do not have access to the items in everyday life, thus removing the possible saturation we experienced with our activities. Furthermore, the tangible items can be small items such as erasers or pencil sharpeners, as demonstrated in the FDP. There are, however, some disadvantages related to using tangible reinforcers. Tangible rewards cost money, and if a future intervention is to be implemented on a large scale, it could be a substantial cost even though the rewards are small, inexpensive items. Furthermore, and maybe of most importance, group activities have more acceptance in the population compared to a physical item as a reward for a child. There are also examples of studies that have managed to obtain good effects with activities as reinforcers, such as PALS. Considering that the children in the lower grades already have many activities, it might be difficult to find activities that will work as reinforcers. However, due to the importance of the acceptance of the teachers and parents, I do recommend activities over tangible rewards if suitable activities are found.

Another explanation for the lack of results could be that the time between the behaviour and end reinforcer was too long and thus that the principle of immediacy was violated. With the end reinforcer delivered on Friday most of the time, the children could have as much as four days between the behaviour and end reinforcer. However, as tokens were delivered close to the behaviour this should not be a major issue. Furthermore, we could not observe an increase in behaviour throughout the week. Although, this may be explained by the large day-to-day variation, which would have made it difficult to observe an increase as we came closer to the end reinforcer. However, the PALS intervention that also uses a token economy system can go several weeks between end reinforcers, which suggest that a four-day delay should not be problematic. If the time between the behaviour and end reinforcer is not a problem, this supports the notion that the suggested reinforcers were not actual reinforcers, or at least that they were outcompeted by other stimuli in the environment. However, PALS is large, complex intervention with more elements than merely a token economy, thus the effect cannot be isolated to the token economy alone (Natvig & Eng, 2012).

5.2.11.2 Modelling

In addition to the possible lack of actual reinforcers, the absence of teacher modelling due to the children watching TV could also contribute to explaining the fruit and vegetable consumption data. With attention being one of the factors needed for observational learning to take place, the TV watching is a problem because it is incompatible with attention toward the teacher. When one of the major behaviour change principles failed to be implemented, it is

expected that the effect of the intervention will also weaken. However, some studies show that watching TV can increase consumption (Mathur & Stevenson, 2015). Therefore, which effects TV watching during lunchbreaks had on fruit and vegetable consumption is unclear.

5.3 Conclusion

The findings in this feasibility study show that simply distributing F&V does not provide a satisfactory level of consumption, at least not for vegetables. Regarding the intervention, the study found that pupils understood and enjoyed taking part in the programme. Furthermore, the intervention had a high completion rate. However, while the study had a high consent rate and suitability for the teachers and that satisfaction amongst the teachers and the education material all met the minimum criteria, there are clear improvements that could and should be made. The study also revealed several unexpected events that should be taken into account when revising the protocol. With mixed effect results and no groups that satisfied the success criteria for either fruit or vegetable consumption, there were no clear indications of effect. Finally, the study found that the eligibility criteria, suggested reinforcers, and delivery of teacher modelling were not feasible. Therefore, in its current state, the intervention is not considered feasible nor is it believed to provide the desired effect on fruit and vegetable consumption. Further improvements are needed for the intervention to be considered feasible and effective, especially regarding the choice of reinforcers and teacher modelling.

5.4 Implications and suggestions for future work

There are several aspect that are necessary to explore before a future study can commence. Of the several points mentioned in this thesis, I will highlight two. One, a mapping of reinforcers should be conducted. A place to start could be a review of which efficient reinforcers are used in school interventions in countries are is comparable to Norway. Furthermore, extensive testing should be conducted to ensure that the reinforcers are efficient in Norwegian schools on a class level.

Another point that needs to be addressed before a study can be developed is the use of digital aids during the lunchbreak. As shown in the FDP, digital aids can be efficient peer models. However, in addition to the consumption of food, the lunchbreak also serves a social function. The current national recommendations for food and meals in schools states: “Schools must ensure that there is suitable options for eating available that also supports the meals social function” (Helsedirektoratet, 2015 p. 17). I would question if the social aspect is accounted for at present. To help guide the development of this and other school interventions, this question needs to be addressed.

Finally, a future school intervention must navigate the ever-changing political suggestions regarding school lunch. Recently the largest party in Norway suggested that schools should start serving a school meal (Aftenposten, 2016). If this suggestion is implemented, other types of interventions might be more feasible. There is, for instance, interesting work being done regarding school meals and nudge that intrudes very little on the remaining school day (Miller et al., 2016).

References

Aftenposten. (2016). Ap vil innføre skolemat for alle. Retrieved from http://www.aftenposten.no/norge/politikk/Ap-vil-innfore-skolemat-for-alle-604267b.html

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, N.J: Prentice Hall.

Bandura, A., Ross, D., & Ross, S.A. (1961). Transmission of aggression through imitation of aggressive models. *Journal of Abnormal and Social Psychology*, *63*, 575–582.

Bandura, A., Ross, D., & Ross, S. A. (1963a). A comparative test of the status envy, social power, and secondary reinforcement theories of identificatory learning. *The Journal of Abnormal and Social Psychology*, *67*(6), 527–534. doi:10.1037/h0046546

Bandura, A., Ross, D., & Ross, S. A. (1963b). Imitation of film-mediated aggressive models. *The Journal of Abnormal and Social Psychology*, *66*(1), 3.

Bere, E., Veierød, M. B., Bjelland, M., & Klepp, K. I. (2006). Outcome and process evaluation of a Norwegian school-randomized fruit and vegetable intervention: Fruits and Vegetables Make the Marks (FVMM). *Health Education Research*, *21*(2), 258–267. doi:10.1093/her/cyh062

Bere, E., Veierød, M. B., & Klepp, K.-I. (2005). The Norwegian School Fruit Programme: evaluating paid vs. no-cost subscriptions. *Preventive Medicine*, *41*(2), 463–470. doi:http://doi.org/10.1016/j.ypmed.2004.11.024

Bjelland, M., Hausken, S. E. S., Bergh, I. H., Grydeland, M., Klepp, K.-I., Andersen, L. F., . . . Lien, N. (2015). Changes in adolescents' and parents' intakes of sugar-sweetened beverages, fruit and vegetables after 20 months: results from the HEIA study – a

comprehensive, multi-component school-based randomized trial. *Food & Nutrition Research*, *59*(1), 25932. doi:10.3402/fnr.v59.25932

Bingham, S., Cassidy, A., Cole, T., Welch, A., Runswick, S., Black, A., . . . Key, T. (1995). Validation of weighed records and other methods of dietary assessment using the 24 h urine nitrogen technique and other biological markers. *British Journal of Nutrition*, *73*(04), 531–550.

Bolarinwa, O. (2015). Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. *Nigerian Postgraduate Medical Journal*, *22*(4), 195–201. doi:10.4103/1117-1936.173959

Border Bia. (unknown). Research & evaluation. Retrieved from http://www.fooddudes.ie/html/research.html

Børge Strømgren Dag Gladmann Sørheim. (2015). Evaluering av the Good Behavior Board Game, en variant av the Good Behavior Game. *Norsk Tidskrift for Afterdsanalyse*, *42*(1), 1–19.

Castro, M. (2007). Placebo versus best-available-therapy control group in clinical trials for pharmacologic therapies: which is better? *Proceedings of the American Thoracic Society*, *4*(7), 570–573. doi:10.1513/pats.200706-073JK

Choi, B. C. K., & Pak, A. W. P. (2005). A catalog of biases in questionnaires. *Preventing Chronic Disease*, *2*(1), A13.

Cooper, J. O., Heron, T. E., & Heward, W. L. (2014). *Applied behavior analysis*. (2nd ed.). Upper Saddle River, NJ: Pearsons Education.

Dauchet, L., Amouyel, P., Hercberg, S., & Dallongeville, J. (2006). Fruit and vegetable consumption and risk of coronary heart disease: a meta-analysis of cohort studies. *The Journal of Nutrition*, *136*(10), 2588–2593.

Evans, C. E., Christian, M. S., Cleghorn, C. L., Greenwood, D. C., & Cade, J. E. (2012). Systematic review and meta-analysis of school-based interventions to improve daily fruit and vegetable intake in children aged 5 to 12 y. *American Journal of Clinical Nutrition, 96*(4), 889–901. doi:10.3945/ajcn.l i 1.030271

Feldrnan, E. B. (2001). Fruits and vegetables and the risk of stroke. *Nutrition Reviews*, *59*(1), 24–27.

Filcheck, H. A., & McNeil, C. B. (2004). The use of token economies in preschool classrooms: practical and philosophical concerns. *Journal of Early and Intensive Behavior Intervention*, *1*(1), 94–104.

Grattan, J., & Demchak, M. A. (2014). Tips for home or school using reinforcement appropriately. Retrieved from http://www.unr.edu/ndsip/tipsheets/27-Guidelines_for_Using_%20Reinforcers.pdf

Grimm, P. (2010). Social desirability bias. In *Wiley international encyclopedia of marketing*. John Wiley & Sons, Ltd.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). Multivariate Data Analysis, Pearson Prentice Hall. *Upper Saddle River, NJ*.

Hansen, L. B., Myhre, J. B., & Andersen, L, F. (2017) *UNGKOST 3 Landsomfattende kostholdsundersøkelse blant 4-åringer i Norge, 2016*. Retrieved from https://www.fhi.no/globalassets/dokumenterfiler/rapporter/rapport-ungkost-3-landsomfattende-kostholdsundersokelse-blant-4-aringer-i-norge-2016.pdf

Hansen, L. B., Myhre, J. B., Johansen, A. M. W., Paulsen M. M., & Andersen, L. F. (2016). *UNGKOST 3 Landsomfattende kostholdsundersøkelse blant elever i 4. -og 8. klasse i Norge, 2015*. Retrieved from https://www.fhi.no/globalassets/dokumenterfiler/rapporter/ungkost-rapport-24.06.16.pdf

Hartley, L., Igbinedion, E., Holmes, J., Flowers, N., Thorogood, M., Clarke, A., . . . Rees, K. (2013). Increased consumption of fruit and vegetables for the primary prevention of cardiovascular diseases. *Cochrane Database of Systematic Reviews* (6). doi:10.1002/14651858.CD009874.pub2

Hawkins, N. G., Sanson-Fisher, R. W., Shakeshaft, A., D'Este, C., & Green, L. W. (2007). The multiple baseline design for evaluating population-based research. *Am J Prev Med, 33*(2), 162–168. doi:http://dx.doi.org/10.1016/j.amepre.2007.03.020

He, F., Nowson, C., Lucas, M., & MacGregor, G. (2007). Increased consumption of fruit and vegetables is related to a reduced risk of coronary heart disease: meta-analysis of cohort studies. *J Hum Hypertens, 21*(9), 717–728.

He, F. J., Nowson, C. A., & MacGregor, G. A. (2006). Fruit and vegetable consumption and stroke: meta-analysis of cohort studies. *The Lancet, 367*(9507), 320–326.

Hebert, J. R., Clemow, L., Pbert, L., Ockene, I. S., & Ockene, J. K. (1995). Social desirability bias in dietary self-report may compromise the validity of dietary intake measures. *International Journal of Epidemiology, 24*(2), 389–398. doi:10.1093/ije/24.2.389

Helsedirektoratet. (2014). Anbefalinger om kosthold, ernæring og fysisk aktivitet. Retrieved from https://helsedirektoratet.no/Lists/Publikasjoner/Attachments/806/Anbefalinger-om-kosthold-ernering-og-fysisk-aktivitet-IS-2170.pdf

Helsedirektoratet. (2015). Nasjonal faglig retningslinje for mat og måltider i skolen Del 1: Barneskole og skolefritidsordning. IS-2371BM Retrieved from https://helsedirektoratet.no/Documents/NFR/skolemat/mat-og-maltider-barneskole-og-SFO-nfr-bokmal.pdf

Helsedirektoratet. (2017). Barnehage – mat, måltider, kosthold: Råd om mat til barn. Retrieved from https://helsedirektoratet.no/folkehelse/kosthold-og-ernering/barnehage-mat-maltider-kosthold#råd-om-mat-til-barn-

Horne, P. J., Hardman, C. A., Lowe, C. F., Tapper, K., Le Noury, J., Madden, P., . . . Doody, M. (2009). Increasing parental provision and children's consumption of lunchbox fruit and vegetables in Ireland: the Food Dudes intervention. *Eur J Clin Nutr, 63*(5), 613–618.

Horne, P. J., Tapper, K., Lowe, C. F., Hardman, C. A., Jackson, M. C., & Woolner, J. (2004). Increasing children's fruit and vegetable consumption: a peer-modelling and rewards-based intervention. *Eur J Clin Nutr, 58*(12), 1649–1660. doi:10.1038/sj.ejcn.1602024

Johnson, R. L., & Morgan, G. B. (2016). *Survey scales: a guide to development, analysis, and reporting*. Guilford Publications.

Joshipura, K. J., Hu, F. B., Manson, J. E., Stampfer, M. J., Rimm, E. B., Speizer, F. E., . . . Willett, W. C. (2001). The effect of fruit and vegetable intake on risk for coronary heart disease. *Annals of Internal Medicine, 134*(12), 1106–1114.

Kreuter, F., Presser, S., & Tourangeau, R. (2008). Social desirability bias in CATI, IVR, and web surveys: the effects of mode and question sensitivity. *Public Opinion Quarterly, 72*(5), 847–865. doi:10.1093/poq/nfn063

Lancaster, G. A. (2015). Pilot and feasibility studies come of age! *Pilot and Feasibility Studies, 1*(1), 1–4. doi:10.1186/2055-5784-1-1

Lee, E. C., Whitehead, A. L., Jacques, R. M., & Julious, S. A. (2014). The statistical interpretation of pilot trials: should significance thresholds be reconsidered? *BMC Medical Research Methodology, 14*(1), 1–8. doi:10.1186/1471-2288-14-41

Leung, L. (2015). Validity, reliability, and generalizability in qualitative research. *Journal of Family Medicine and Primary Care, 4*(3), 324–327. doi:10.4103/2249-4863.161306

Lowe, F., Horne, P., Viktor, S., Kelly, P., Pears, S., & Anthony, T. (2011). Case study: Food Dudes healthy eating programme: learning to like fruit and vegetables. In J. French,

R. Merritt, & L. Reynolds (Eds.), Social marketing case book (pp. 219–231). Sage Publications

Marmot, M., Atinmo, T., Byers, T., Chen, J., Hirohata, T., Jackson, A., . . . Mann, J. (2007). Food, nutrition, physical activity, and the prevention of cancer: a global perspective. Retrieved from http://www.wcrf.org/sites/default/files/Second-Expert-Report.pdf

Mathur, U., & Stevenson, R. J. (2015). Television and eating: repetition enhances food intake. *Frontiers in Psychology, 6*, 1657. doi:10.3389/fpsyg.2015.01657

Miller, G. F., Gupta, S., Kropp, J. D., Grogan, K. A., & Mathews, A. (2016). The effects of pre-ordering and behavioral nudges on National School Lunch Program participants' food item selection. *Journal of Economic Psychology*, *55*4–16.

Miller, L. K. (2006). *Principles of everyday behavior analysis*. Thomson/Wadsworth.

Mytton, O. T., Nnoaham, K., Eyles, H., Scarborough, P., & Mhurchu, C. N. (2014). Systematic review and meta-analysis of the effect of increased vegetable and fruit consumption on body weight and energy intake. *BMC Public Health, 14*(1), 886.

Natvig, H., & Eng, H. (2012). Beskrivelse og vurdering av tiltaket PALS – Positiv atferd, støttende læringsmiljø og samhandling. Retrieved from https://www.ungsinn.no/post_tiltak/pals-positiv-atferd-stottende-laeringsmiljo-og-samhandling-2/

O'Cathain, A., Hoddinott, P., Lewin, S., Thomas, K. J., Young, B., Adamson, J., . . . Donovan, J. L. (2015). Maximising the impact of qualitative research in feasibility studies for randomised controlled trials: guidance for researchers. *Pilot and Feasibility Studies, 1*(1), 32. doi:10.1186/s40814-015-0026-y

Orsmond, G. I., & Cohn, E. S. (2015). The distinctive features of a feasibility study objectives and guiding questions. *OTJR: Occupation, Participation and Health, 35*(3), 169–177.

Perisic, I., & Rosner, B. (1999). Comparisons of measures of interclass correlations: the general case of unequal group size. *Statistics in Medicine, 18*(12), 1451–1466

Powell, P., Spears, K., & Rebori, M. (2010). What is obesogenic environment?. Retrieved from https://www.unce.unr.edu/publications/files/hn/2010/fs1011.pdf

Presti, G., Cau, S., Oppo, A., & Moderato, P. (2015). Increased classroom consumption of home-provided fruits and vegetables for normal and overweight children: results of the Food Dudes program in Italy. *Journal of Nutrition Education & Behavior, 47*(4), 338–344. doi:10.1016/j.jneb.2015.04.331

Pringle, B. A., Colpe, L. J., Blumberg, S. J., Avila, R. M., & Kogan, M. D. (2012). Diagnostic history and treatment of school-aged children with autism spectrum disorder and special health care needs. *NCHS Data Brief, 97*, 1–8.

Psychology Dictionary. (unkown). What is MODELING. Retrieved from http://psychologydictionary.org/modeling/

Rhoda, D. A., Murray, D. M., Andridge, R. R., Pennell, M. L., & Hade, E. M. (2011). Studies with staggered starts: multiple baseline designs and group-randomized trials. *American Journal of Public Health, 101*(11), 2164–2169. doi:10.2105/AJPH.2011.300264

Skinner, B. F. (1953). Science and human behavior. New York: The Free Press.

Skinner, B. F. (1969). Contingencies of reinforcement: a theoretical analysis. New York: Appleton-Century-Crofts

Sosial – og Helsedirektoratet. (2005). *Frukt og grønnsaker i skolen Beregning av samfunnsøkonomisk lønnsomhet* (IS-1281). Retieved from https://helsedirektoratet.no/Lists/Publikasjoner/Attachments/64/Frukt-og-gronnsaker-i-skolen-beregning-av-samfunnsokonomisk-lonnsomhet-IS-1281.pdf

Sørlie, M. A., Ogden, T., Arnesen, A., Olseth, A. R., & Hansen, W. M. (2014). Skoleledere om PALS. Retrieved from https://www.utdanningsnytt.no/globalassets/filer/pdf-av-spesialpedagogikk/2014/spesialpedagogikk--1-2014.pdf

te Velde, S. J., Brug, J., Wind, M., Hildonen, C., Bjelland, M., Pèrez-Rodrigo, C., & Klepp, K. I. (2008). Effects of a comprehensive fruit- and vegetable-promoting school-based intervention in three European countries: the Pro Children study. *British Journal of Nutrition*, *99*, 893–903. doi: 10.1017/S000711450782513X

Thabane, L., Ma, J., Chu, R., Cheng, J., Ismaila, A., & Rios, L. P. (2010). A tutorial on pilot studies: the what, why and how. *BMC Med Res Methodol*, *10*. doi:10.1186/1471-2288-10-1

Totland, T. H., Gebremariam, M. K., Lien, N., Bjelland, M., Grydeland, M., Bergh, I. H., . . . Andersen, L. F. (2013). Does tracking of dietary behaviours differ by parental education in children during the transition into adolescence? *Public Health Nutrition, 16*(04), 673–682.

Totland, T. H., Melnæs, B. K., Lundberg-Hallén, N., Helland-Kigen, K. M., Lund-Blix, N. A., Myhre, J. B., . . . Andersen, L. F. (2012). Norkost 3 En landsomfattende kostholdsundersøkelse blant menn og kvinner i Norge i alderen 18–70 år, 2010–11. Retrieved from https://www.lhl.no/globalassets/dokumenter/norkost-3-is-20001.pdf

Trochim, W. M. K. (2006). Measurement validity types. Retrieved from https://www.socialresearchmethods.net/kb/measval.php

Upton, D., Upton, P., & Taylor, C. (2013). Increasing children's lunchtime consumption of fruit and vegetables: an evaluation of the Food Dudes programme. *Public Health Nutrition, 16*(6), 1066–1072. doi:http://dx.doi.org/10.1017/S1368980012004612

U.S. Department of Health and Human Services and U.S. Department of Agriculture. (2015). *2015–2020 dietary guidelines for Americans*. (8th ed.). Retrieved from https://health.gov/dietaryguidelines/2015/resources/2015-2020_Dietary_Guidelines.pdf

Van de Mortel, T. F. (2008). Faking It: social desirability response bias in self-report research. *Australian Journal of Advanced Nursing*, *25*(4), 40.

Wang, X., Ouyang, Y., Liu, J., Zhu, M., Zhao, G., Bao, W., & Hu, F. B. (2014). Fruit and vegetable consumption and mortality from all causes, cardiovascular disease, and cancer: systematic review and dose-response meta-analysis of prospective cohort studies. *BMJ*, *349*, g4490.

Wengreen, H., Joyner, D., & Madden, G. (2015). FITGAME: A school-wide game-based fruit and vegetable intervention. *Journal of Nutrition Education & Behavior*, *47*(4), S74–S74. doi:10.1016/j.jneb.2015.04.196

Wengreen, H. J., Madden, G. J., Aguilar, S. S., Smits, R. R., & Jones, B. A. (2013). Incentivizing children's fruit and vegetable consumption: Results of a United States pilot study of the Food Dudes program. *Journal of Nutrition Education & Behavior*, *45*(1), 54–59. doi:10.1016/j.jneb.2012.06.001

World medical association (2013). WMA declaration of Helsinki – ethical principles for medical research involving human subjects. Retrieved from https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/

Appendix 1

Opplæringsverktøy for lærere i atferdsintervensjonen Frukt og Grønt i Skolen

Innhold

Kort om prosjektet 82
Baseline fase 82
Intervensjonsfasen 83
Læreraktivitet i klasserommet 84
Kriterier 85
Kriteriet for liten stjerne (stjerne i individuelt skjema) 85
Kriteriet for aktivitet (antall stjerner på klasseskjema) 85
Aktiviteter 86
Hva om …? 86
Et barn mister en frukt eller grønnsak på gulvet? 86
Barna spør hvorfor studentene er i klasserommet? 86
Barna vil beholde sine egne skjemaer som er fylt ut? 86
Barna ikke ønsker å stå foran klassen å si hva de har spist? 86

Kort om prosjektet

Frukt og grønt i skolen er en intervensjon som har som hensikt å øke frukt- og grønnsakinntaket blant skolebarn. Prosjektet vil benytte modellærings- og belønningsprinsipper for å påvirke barnas frukt- og grønnsaksinntak. Prosjektet består av to faser: en baseline fase der barna får utdelt frukt og grønnsaker og en intervensjonsfase, der selve tiltaket og modelleringen foregår.
En praksisstudent fra HiOA vil være tilstede på skolen under prosjektet for å gjøre forarbeid, etterarbeid og bidra til gjennomføringen.
Lærerens rolle i løpet av prosjektet er først og fremst å administrere belønningssystemet og legge til rette for god modellæring som beskrevet i dette dokumentet.

Baseline fase

Utstyr: Personlige matbokser (utlevert av HiOA).
I baseline fasen vil barna få utdelt frukt og grønnsaker ved starten av matpausen. Det gjøres ikke tiltak overfor barna utover dette, men det er viktig at følgende regler blir gitt på en enkel måte som barna forstår:

- Barna kan spise så mye frukt og/eller grønnsaker de selv ønsker.
- Restene som de ikke ønsker/rekker å spise skal de la ligge i matboksen.

- De får ikke ta med rester av frukt eller grønnsaker ut i friminuttet.
- De får ikke lov til å dele frukt og grønnsaker med hverandre.

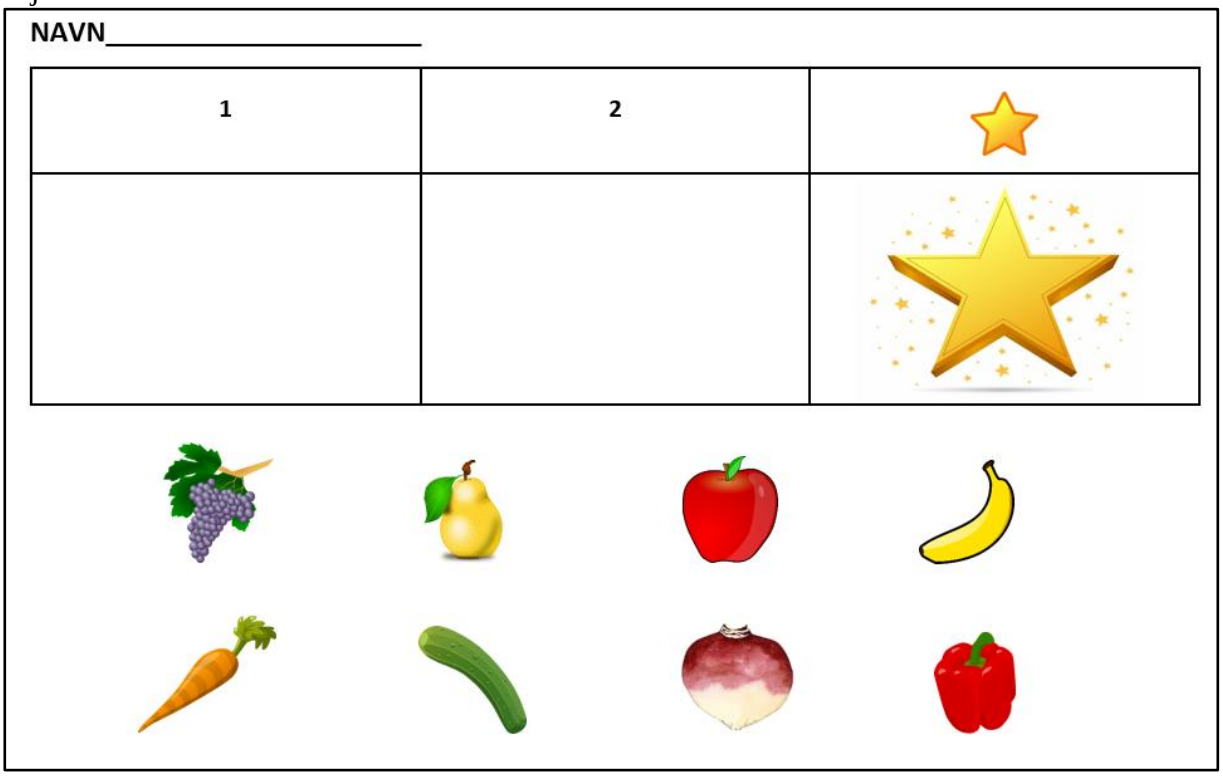
I baseline fasen vil du som lærer oppføre deg slik du pleier. Hvis du vanligvis er aktiv når det gjelder ros av elever som spiser frukt eller grønnsaker, er det ønskelig at ros holdes på et moderat nivå i baseline fasen. Ros/vise interesse for barn som spiser frukt og grønt er en del av intervensjonen. Det må være mulig for deg å kunne øke frekvensen på ros/vist interesse fra baseline fase til intervensjonsfase.

Intervensjonsfasen

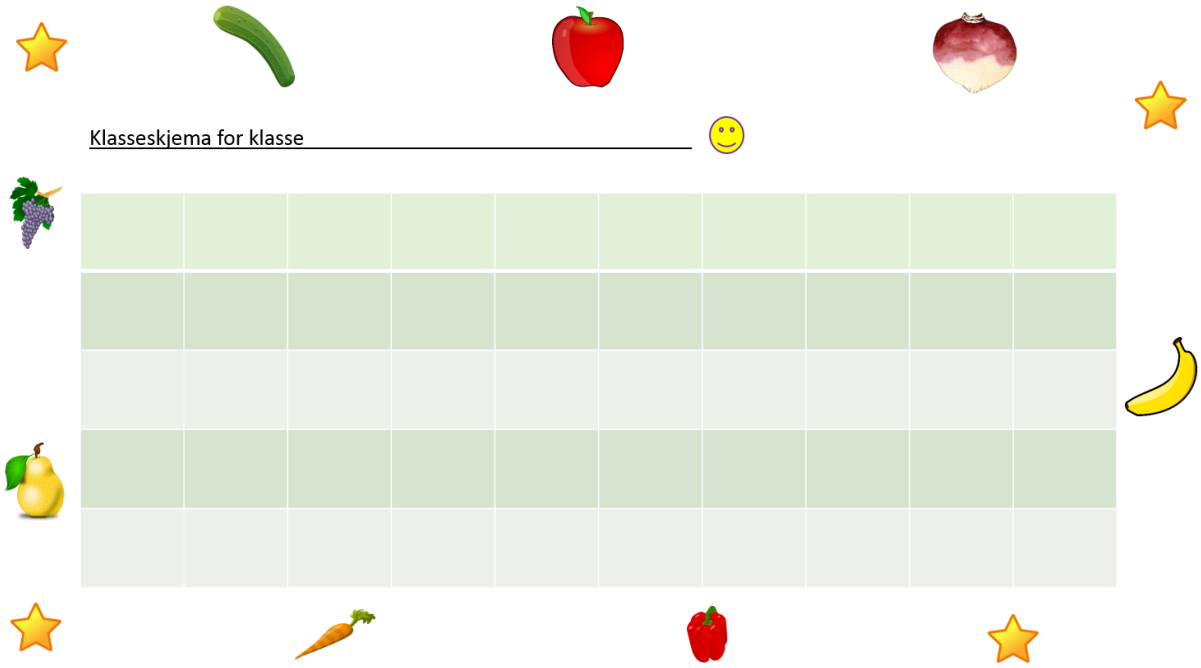
Utstyr:

- Hvert barn vil få utdelt et personlig skjema.
- I tillegg får klassen utdelt et felles skjema som skal henges opp synlig i klasserommet.
- En aktivitetsboks blir utdelt til klassen I boksen er det lapper med aktiviteter som barna liker å gjøre.

Når intervensjonsfasen starter, er det viktig at barna blir informert om at reglene fra baseline fasen fortsatt gjelder.
I tillegg må de få vite at hvis de spiser frukt og grønnsaker tilsvarende kriteriet, (kriteriet vil variere og er spesifisert under kriterier) vil de få en stjerne i skjemaet sitt. Skjemaet til barna vil likne på skjema presentert under. De vil få et nytt skjema hver gang de har fylt ut et skjema.



Når et barn har fått to stjerner i skjemaet sitt, vil det få en stor stjerne som plasseres på klasseskjemaet som henger synlig i klasserommet. Denne stjernen skal barna henge opp selv. De skal da samtidig si høyt hvilken frukt eller grønnsak de likte best. For å hjelpe barna å huske hva de har spist er det bilder av alle fruktene og grønnsakene i barnas individuelle skjema. Et tips kan være å du eller de setter en ring rundt den frukten og grønnsakene de spiste når de fikk en stjerne. Klasseskjemaet vil likne på skjemaet presentert under.



Når klassen i fellesskap har samlet et gitt antall klistremerker (antall spesifisert under Kriterier) på klasseskjemaet, vil du trekke en aktivitet fra aktivitetsboksen.
Ved intervensjonsstart vil det også være naturlig å informere barna om aktivitetene de kan oppnå ved å spise frukt og grønnsaker (aktivitetene er listet opp under Aktiviteter). For å være sikker på at aktivitetene faller i smak, bør barna få muligheten til å påvirke hvilke aktiviteter som skal med. Denne utvelgelsesprosessen kan f. eks. gjennomføres ved handsopprekning.
Barna må også informeres om at de skal si høyt til klassen hvilken frukt og grønnsak i skjemaet de likte best etter at de har klistret en stor stjerne på klasseskjemaet. Dette er noe du kan minne dem på når de får utlevert stor stjerne.
Informasjon til barna kort oppsummert:

- Barna må forstå at det å spise en gitt mengde frukt og grønnsaker vil lede til en stjerne i sitt individuelle skjema.
- To stjerner i deres individuelle skjema vil gi dem en stor stjerne i klasseskjema.
- Barna skal si høyt hvilken frukt/grønnsak de likte best når de klistrer på stor stjerne på klasseskjema
- Når de oppnår et gitt antall stjerner i klasseskjema, vil læreren trekke en aktivitet fra aktivitetsboksen.

Læreraktivitet i klasserommet

I intervensjonsfasen vil du ha noen enkle oppgaver:

- Gi ros/oppmerksomhet til elever som spiser/spiser opp frukt og grønnsaker
- Vurdere om barnet har oppnådd kriteriet for klistremerke (vurderes visuelt, se Kriterier)
- Levere ut små og store klistremerker til barna
- Trekke aktivitet fra aktivitetsboksen og gjennomføre aktiviteten med elevene
- Spis frukt og grønnsaker foran barna (du vil få utdelt frukt og grønnsaker i intervensjonsperioden)

- Minne barna om at de som får en stjerne til klasseskjemaet skal si hva som er favorittfrukten og favorittgrønnsaken deres.

Modellæring har to aspekter: 1) lærer og student vil fungere som modeller for elevene, og 2) eleven vil opptre som modeller for hverandre.
Hvis du opplever at alle barna har den samme favorittfrukten -eller grønnsaken, kan du variere hva du ber dem om å si til klassen. F.eks. istedenfor hva som er favorittfrukten kan du be barnet si hvilke frukter og grønnsaker han eller hun spiste for å få den stjernen.
Du kan også trekke frem elever som spiser frukt og grønnsaker andre vegrer seg for å spise. Hvis mange elever sliter med å spise paprika, kan du spørre et barn du vet har spist paprika om hva hun eller han har spist. Deretter kan du rose eleven for å ha spist paprika foran klassen.
Dette vil kunne har flere effekter:

1. Sjansen for at barnet spiser paprika igjen vil øke, fordi det opplevde noe positivt knyttet til det.
2. Eleven fungerer som en god modell for paprikaspising for de andre barna.
3. De andre barna ser at en klassekamerat fikk ros av en viktig person fordi han/hun spiste paprika.

Kriterier
Kriteriene for liten stjerne og aktivitet vil variere under intervensjonsfasen.

Kriteriet for liten stjerne (stjerne i individuelt skjema)

Intervensjonsuke 1: Barna vil motta en stjerne hvis de spiser opp halvparten av frukten og halvparten av grønnsakene. Om barnet har oppnådd kriteriet vurderes visuelt.

Intervensjonsuke 2 og utover: Barna vil motta en stjerne for å spise opp både frukten og grønnsakene.
Mengdene frukt og grønnsaker kan forandres utover i intervensjonen hvis mengden viser seg å være for liten eller for stor. Hvis dette skulle skje, vil studenten på skolen si ifra til meg, så jeg kan justere mengden som leveres.

Kriteriet for aktivitet (antall stjerner på klasseskjema)

Intervensjonsuke 1: Den første uken vil kriteriet for aktivitet settes lavt. Denne uken vil det være lagt opp til to muligheter for aktiviteter (f. eks. tirsdag og fredag). Klasseskjemaet må ha halvparten så mange stjerner som det er elever i klassen (en klasse med 30 elever må ha 15 stjerner).
Tips: Det anbefales at aktivitetene som benyttes første uken er enkle å gjennomføre (f. eks. tegning/fargelegging, høytlesning) slik at det ikke blir for krevende med to aktiviteter på en uke.

Intervensjonsuke 2 og 3: I uke 2 og 3 skal barna få muligheten til å oppnå aktivitet én gang per uke. Antall stjerner vil da settes til elevantallet i klassen +1. (en klasse på 30 elever vil ha 31 stjerner som krav).

Intervensjonsuker etter uke 3: Kravet til stjerner i klasseskjema etter uke 3 vil vurderes fra uke til uke. Kravet baseres på tidligere ukers oppnåelser. Hvis f. eks elevene oppnådde 35 stjerner i uke 3, når kravet var 30, kan kravet økes til 35-40. Her må du gjøre en vurdering sammen med praksisstudenten.

Aktiviteter

Aktiviteter som kan være passende å bruke kan være:

- Høytlesning
- Tegning og fargelegging
- Leketime i gymsalen eller annet passende rom
- Ekstra leketid ute
- Spilletime, f. eks. brettspill eller andre spill hvis skolen har det
- En episode eller deler av en episode f. eks. Newton, NRK super tegnefilm eller andre programmer som barna liker
- Sang og dans til video eller CD
- Leketime med leke tatt med hjemmefra

Hvis du tenker det er andre aktiviteter som er aktuelle, må du gjerne inkludere dem.
Hvilke av aktivitetene som inkluderes i aktivitetsboksen er avhengig av hva du vurderer er gjennomførbart som elevene liker.
Aktivitetsboksen som aktivitetene trekkes fra vil lages på en slik måte at du kan trekke spesifikke lapper om det er nødvendig (f. eks. uker der du har reservert gymsalen, eller det har kommet snø og barna har satt opp «Leke i snøen» som aktivitet).
Lengden på aktivitetene kan du vurdere, og det er ikke noe problem om de forskjellige aktivitetene varierer i lengde. Det er du som kjenner dine elever og skoletimer best.

Hva om …?
Et barn mister en frukt eller grønnsak på gulvet?
Hvis et barn mister en frukt- eller grønnsaksbit på bakken, vil det stå en boks i klasserommet som barnet kan få en ny bit av. Den nye biten tilpasses slik at den er ca. like stor som biten som ble mistet (dette ordner studenten).

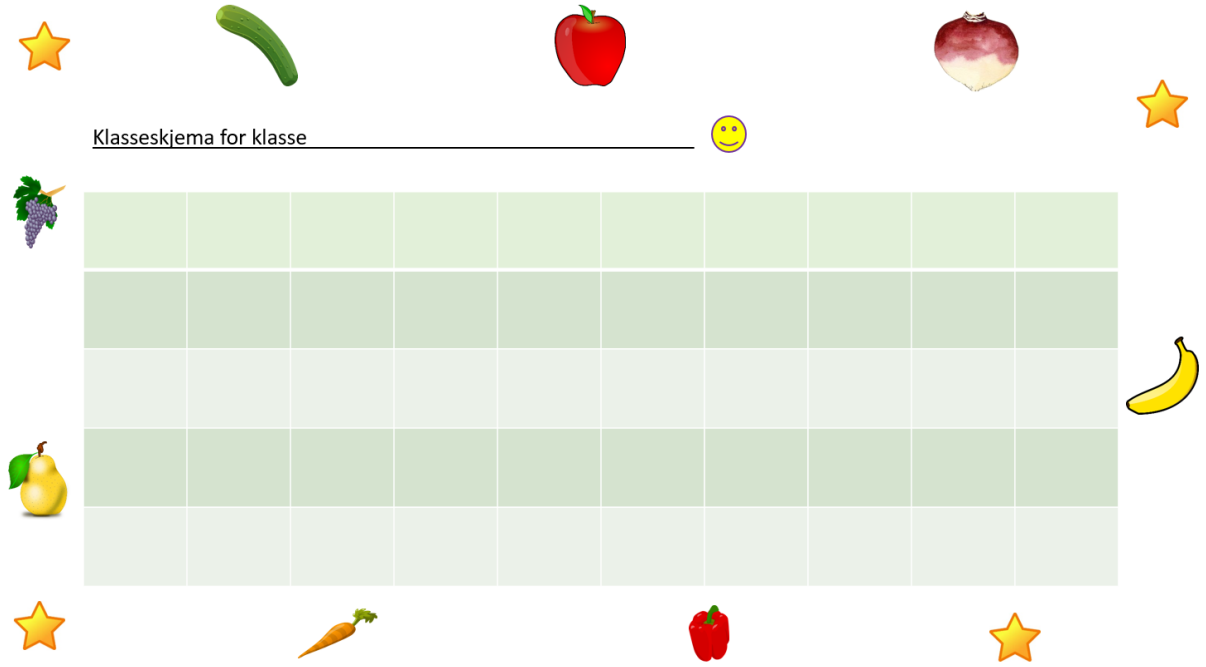
Barna spør hvorfor studentene er i klasserommet?
Det er viktig at du ikke forteller barna at de blir målt. Dette er for å sikre at de ikke endrer atferd basert på informasjonen. Du kan heller si at studentene også går på skole og at de skal hjelpe til med å dele ut frukt og grønnsaker.

Barna vil beholde sine egne skjemaer som er fylt ut?
I utgangspunktet anbefaler vi at barna ikke får beholde skjemaer som er ferdig utfylt under intervensjonen. Dette vil kunne skape problemer med å skille mellom hva som er nye og gamle skjemaer. Et alternativ kan være å samle inn alle skjemaene som er ferdig utfylt og levere de ut igjen til elevene ved slutten av intervensjonen så de kan få ta de med hjem om de ønsker det.
Det anbefales også at de individuelle skjemaene barna får utlever som ikke er ferdig utfylt beholdes i klasserommet. Hvis alle barna skal ta vare på disse i egen sekk vil det fort kunne oppstå uheldige situasjoner der noen glemmer eller mister et påbegynt skjema.

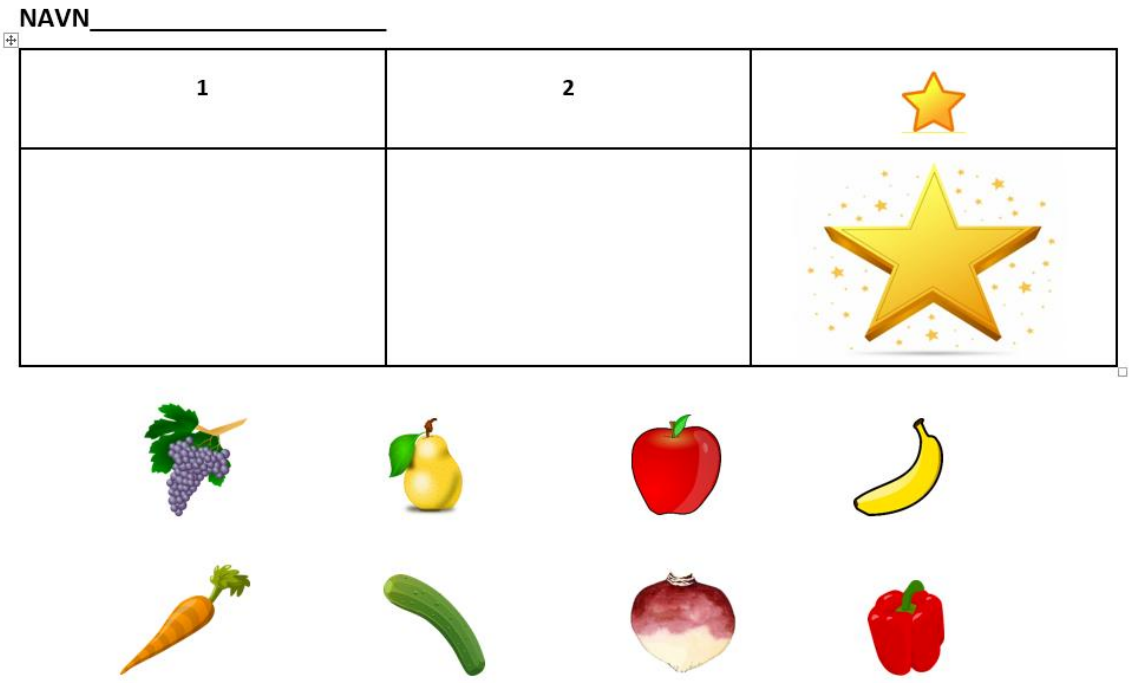
Barna ikke ønsker å stå foran klassen å si hva de har spist?
Hvis det er noen barn som er litt sjenerte og ikke ønsker å snakke foran klassen kan de selvfølgelig slippe det. Dette styrer du som du pleier i slike situasjoner.

Appendix 2

Class token economy board



Individual token economy board



Appendix 3

Spørreskjema for prosjektet frukt og grønt i skolen.

Når du svarer på spørreskjemaet skal du svare på dine egne vegne og ikke på vegne av skoleansatte som gruppe. Vennligst svar på alle holdningsutsagnene i spørreskjemaet og benytt kommentarfeltene hvis det er noe du ønsker å tilføye.

Når du svarer på holdningsutsagnene skal du svare på om du mener skolen kunne gjennomført intervensjonen uten hjelp utenfra, altså uten en student på skolen. Intervensjonsoppgavene er følgende: 1) distribusjon av frukt og grønnsaker til barna i klasserommet; 2) spise frukt og grønnsaker foran barna; 3) levere ut små og store stjerner; 4) administrere stjerneskjemaene og 5) levere ut belønning.

Holdningsutsagn	**Sterkt uenig**					**Sterkt enig**
Jeg opplevde at intervensjonen var gjennomførbar.	1	2	3	4	5	6
Jeg erfarte at tidsbruken i intervensjonen var akseptabel.	1	2	3	4	5	6
Jeg mener at arbeidsoppgavene i intervensjonen var overkommelige.	1	2	3	4	5	6
Jeg oppfattet intervensjonen som krevende, med tanke på skolens tilgjengelige ressurser.	1	2	3	4	5	6
Jeg mener at intervensjonen tok for mye tid av skoledagen.	1	2	3	4	5	6
Jeg mener at intervensjonen må forenkles for å være gjennomførbar i skolehverdagen.	1	2	3	4	5	6
Andre kommentarer du har angående din erfaring med gjennomføringen?						

	Sterkt uenig					Sterkt enig
Jeg opplevde at elevene hadde tid til å spise frukten og grønnsakene i lunsjpausen.	1	2	3	4	5	6
Jeg mener utvidet spisetid hadde økt barnas konsum av frukt og grønnsaker.	1	2	3	4	5	6
Jeg mener tid var en begrensende faktor for at barna kunne spise opp frukt og grønnsakene.	1	2	3	4	5	6
Andre kommentarer du har til elevens bruk av tid til å spise frukten og grønnsakene?						

	Sterkt uenig					Sterkt enig
Jeg mener mengden frukten og grønnsakene elevene fikk utlevert i matboksene var passende.	1	2	3	4	5	6
Jeg ville ha økt mengden frukt barna fikk utlevert hvis jeg fikk muligheten.	1	2	3	4	5	6
Jeg ville ha økt mengden grønnsaker barna fikk utlevert hvis jeg fikk muligheten.	1	2	3	4	5	6
Jeg ville ha redusert mengden frukt barna fikk utlevert hvis jeg fikk muligheten.	1	2	3	4	5	6
Jeg ville ha redusert mengdene grønnsaker barna fikk utlevert hvis jeg fikk muligheten.	1	2	3	4	5	6
Andre kommentarer du har angående mengde frukt og grønt?						

	Sterkt uenig					Sterkt enig
Jeg opplevde at barna forsto tegnøkonomi systemet (stjerneskjemaene)	1	2	3	4	5	6
Jeg erfarte at barna forsto at frukt og grønnsaksspising ledet til belønning.	1	2	3	4	5	6
Andre kommentarer du har angående barnas forståelse av programmet?						

	Sterkt uenig					Sterkt enig
Jeg opplevde at modellæringen (at du spiste frukt og grønt foran barna) førte til at barna spiste mer frukt og grønnsaker.	1	2	3	4	5	6
Jeg opplevde ikke at modellæringen (at du spiste frukt og grønt foran barna) påvirket barnas frukt og grønt konsum i noen stor grad.	1	2	3	4	5	6
Jeg opplevde at tegnøkonomiskjemaene (stjerneskjemaene) påvirket barnas frukt og grønt inntak.	1	2	3	4	5	6

Jeg opplevde at belønningen barna kunne oppnå påvirket barnas frukt og grønt inntak.	1	2	3	4	5	6
Jeg hadde takket ja til å delta i prosjektet om jeg ble spurt igjen.	1	2	3	4	5	6
Jeg ville anbefalt programmet til andre skoler.	1	2	3	4	5	6
Jeg erfarte at barna viste interesse for programmet.	1	2	3	4	5	6
Jeg opplever at barna likte å delta i programmet.	1	2	3	4	5	6
Andre kommentarer du har til barnas deltakelse i programmet?						

	Sterkt uenig					Sterkt enig
Jeg leste nøye igjennom opplæringsverktøyet for intervensjonen begynte.	1	2	3	4	5	6
Jeg opplevde at jeg var usikker på hvordan jeg skulle gjennomføre forskjellige arbeidsoppgaver.	1	2	3	4	5	6
Jeg følte at jeg hadde god forståelse av programmet som helhet.	1	2	3	4	5	6
Jeg erfarte at det oppstod situasjoner der jeg ikke visste hvordan jeg skulle gå fram?	1	2	3	4	5	6
Jeg har hatt en klar forståelse av mine arbeidsoppgaver i prosjektet	1	2	3	4	5	6
Jeg forsto informasjonen som var gitt i opplæringsverktøyet.	1	2	3	4	5	6
Jeg syntes informasjonsverktøyet var uklart.	1	2	3	4	5	6
Jeg opplevde kvaliteten på opplæringsverktøyet som god nok til at jeg skulle være i stand til å gjennomføre intervensjonen.	1	2	3	4	5	6
Andre kommentarer som du har til opplæringsverktøyet?						

Hvis du har noen flere kommentarer du ønsker å meddele, for eksempel:

Hva syntes du funket bra?

Hva syntes du funket dårlig?

Andre ting du vil nevne?

Vennligst se over om du har svart på alle holdningsutsagnene.
Takk for at du tok det tid til å svare på denne spørreundersøkelsen.

Appendix 4

Forespørsel om deltakelse i forskningsprosjektet

Frukt og Grønt i Skolen

Bakgrunn og formål

Formålet med studien er å teste gjennomførbarheten av intervensjonen Frukt og Grønt i Skolen. Frukt og Grønt i Skolen er en atferdsintervensjon som har som hensikt å øke skolebarns inntak av frukt og grønnsaker. Som en gjennomførbarhets studie vil denne studien utforske hvordan barna, lærerne og dere selv (foreldre/verger) oppfatter og erfarer deltakelse i studien. Studien vil også utforske praktiske hensyn ved implementering og gjennomføring av intervensjonen, samt utforske eventuelle korttidstrender på inntaket av frukt og grønt.
Prosjektet er en masterstudie ved Høgskolen i Oslo og Akershus og er et samarbeid mellom Institutt for sykepleie og helsefremmendearbeid og Institutt for atferdsvitenskap.
Du/dere og deres barn blir spurt om å delta i studien fordi skolen ditt barn går på har takket ja til å delta i studien.

Hva innebærer deltakelse i studien?

For ditt/deres barn vil deltakelse i studien bety at han/hun vil få utdelt frukt og grønnsaker på skolen hver dag i 6 uker (7. november – 16. desember). Barn som smaker på/spiser opp frukten og grønnsakene vil få et klistremerke som de klistrer på en felles tavle i klasserommet. Ved et gitt antall klistremerker vil barna få en felles belønning som for eksempel 30 minutter ekstra friminutt eller høytlesning. Barna sammen med lærerne og bachelorstudenter vil også fungere som gode rollemodeller for hverandre.
Som forelder vil deres deltakelse i studien bety at dere får et spørreskjema sent hjem der dere oppgir litt bakgrunnsinformasjon (sosioøkonomiske forhold) samt eventuelle overfølsomhetsreaksjoner barnet har mot frukt og grønnsaker som benyttes i prosjektet. Dere vil også motta et spørreskjema ved studieslutt dere dere svarer på hvordan dere som foreldre opplevde og erfarte at barnet deres deltok i studien.

Hva skjer med informasjonen om ditt/dere barn og deg?

Alle personopplysninger vil bli behandlet konfidensielt. De som vil ha tilgang på, eller komme i kontakt med, personopplysningene om deres barn og dere selv vil være masterstudenten som er ansvarlig for prosjektet, veileder og biveileder, samt bachelorstudenter som vil bidra i implementeringen av intervensjonen. All personlig informasjon vil bli anonymisert under intervensjonen. Intervensjonsdata og navnelister vil lagres nedlåst, adskilt fra hverandre.
Prosjektet skal etter planen avsluttes 20. desember. Etter plottingen av dataene er ferdig og kontrollert, vil alle dokumenter som inneholder sensitive opplysninger som for eksempel navnelister og eventuelle overfølsomheter, bli destruert. Den endelige destrueringen av de sensitive opplysningene vil finne sted innen 28. februar 2017.

Frivillig deltakelse

Det er frivillig å delta i studien, og du kan når som helst trekke ditt samtykke uten å oppgi noen grunn. Dersom du trekker deg, vil alle opplysninger om deg bli anonymisert.

Dersom du ønsker å delta eller har spørsmål til studien, ta kontakt med prosjektansvarlig Magnus Haakens på tlf. nr. 41 31 45 11 eller via e-post S185527@stud.hioa.no eller veileder professor Kjell Sverre Pettersen på tlf. nr. 92 04 78 67 e-post kjellsverre.pettersen@hioa

Studien er meldt til Personvernombudet for forskning, NSD - Norsk senter for forskningsdata AS.

Samtykke til deltakelse i studien

Jeg har mottatt informasjon om studien, og er villig til å delta……☐

Jeg samtykker til at mitt barn kan delta i studien……☐

Navn på barn (fornavn og etternavn i blokkbokstaver) _____

(Signert av prosjektdeltaker/forelder/verge, dato)

Appendix 5

Forespørsel om deltakelse i forskningsprosjektet

Frukt og Grønt i Skolen

Bakgrunn og formål

Formålet med studien er å teste gjennomførbarheten av intervensjonen Frukt og Grønt i Skolen. Frukt og Grønt i Skolen er en atferdsintervensjon som har som hensikt å øke skolebarns inntak av frukt og grønnsaker. Som en gjennomførbarhets studie vil denne studien utforske hvordan barna, lærerne og foreldre/verger oppfatter og erfarer deltakelse i studien. Studien vil også utforske praktiske hensyn ved implementering og gjennomføring av intervensjonen, samt utforske eventuelle korttidseffekter.
Prosjektet er en masterstudie ved Høgskolen i Oslo og Akershus og er et samarbeid mellom Institutt for sykepleie og helsefremmendearbeid og Institutt for atferdsvitenskap.

Hva innebærer deltakelse i studien?

For deg som lærer vil deltakelse i studien innebære at du vil delta i implementeringen av prosjektet frukt og grønt i skolen. Under prosjektperioden vil en student være tilstede på skolen å bidra til implementeringen av prosjektet. Etter prosjektslutt vil du motta et spørreskjema som f. eks. vil utforske praktiske forhold ved gjennomføringen, erfaringer etc. Studentene vil også samle data igjennom observasjon under prosjektperioden.

Hva skjer med informasjonen om deg?

Alle personopplysninger vil bli behandlet konfidensielt. De som vil ha tilgang på, eller komme i kontakt med deres opplysninger vil være masterstudenten som er ansvarlig for prosjektet, veileder og biveileder, samt bachelorstudenter som vil bidra i implementeringen av intervensjonen. All personlig informasjon vil bli anonymisert under intervensjonen. Intervensjonsdata og navnelister vil lagres nedlåst, adskilt fra hverandre.
I fremtidige rapporter og publikasjoner som tar utgangspunkt i data fra dette prosjektet vil informasjonen presenteres på en slik måte at deres anonymitet ivaretas.
Prosjektet skal etter planen avsluttes 16. desember. Etter plottingen av dataene er ferdig og kontrollert, vil alle dokumenter som inneholder sensitive opplysninger som for eksempel navnelister bli destruert. Den endelige destrueringen av de sensitive opplysningene vil finne sted innen 28. februar 2017.

Frivillig deltakelse

Det er frivillig å delta i studien, og du kan når som helst trekke ditt samtykke uten å oppgi noen grunn. Dersom du trekker deg, vil alle opplysninger om deg bli anonymisert.

Dersom du ønsker å delta eller har spørsmål til studien, ta kontakt med prosjektansvarlig Magnus Haakens på tlf. nr. 41 31 45 11 eller via e-post S185527@stud.hioa.no eller veileder professor Kjell Sverre Pettersen på tlf. nr. 92 04 78 67 e-post kjellsverre.pettersen@hioa

Studien er meldt til Personvernombudet for forskning, NSD - Norsk senter for forskningsdata AS.

Samtykke til deltakelse i studien

Jeg har mottatt informasjon om studien, og er villig til å delta

---

(Signert av prosjektdeltaker, dato)

Appendix 6

Spørsmål om allergier og overfølsomheter

For å sikre at ingen barn får i seg noe de ikke tåler, samler vi informasjon om eventuelle allergier eller overfølsomheter mot frukt og grønnsaker som vil bli brukt i studien. Hvis ditt/deres barn er allergisk mot en av matvarene som vil bli brukt, vil barnet få erstattet matvaren han/hun ikke tåler med en alternativ matvare.
Vennligst kryss av for frukten(e) og/eller grønnsaken(e) barnet ditt/deres har en allergi eller overfølsomhet mot. Hvis ditt/deres barn verken har allergi eller overfølsomhet, vennligst kryss av for Ingen allergier eller overfølsomheter for frukten og grønnsakene nevnt over.

Frukt

Epler…… ☐
Druer…… ☐
Banan…… ☐
Pære…… ☐

Grønnsaker

Gulrot…… ☐
Kålrot…… ☐
Agurk …… ☐
Paprika …… ☐

Ingen allergier eller overfølsomheter for frukten og grønnsakene nevnt over …… ☐

Navn på barn (fornavn og etternavn i blokkbokstaver)

_____

(Signert av forelder/verge, dato)

Appendix 7
Recommendation from NSD

Sverre Pettersen
Institutt for sykepleie og helsefremmende arbeid Høgskolen i Oslo og Akershus
Postboks 4 St. Olavs plass
0130 OSLO

Vår dato: 06.10.2016 Vår ref: 49614 / 3 / STM Deres dato: Deres ref:

TILBAKEMELDING PÅ MELDING OM BEHANDLING AV PERSONOPPLYSNINGER

Vi viser til melding om behandling av personopplysninger, mottatt 26.08.2016. Meldingen gjelder prosjektet:

49614	*Frukt og grønt i skolen: en atferdsintervensjon*
Behandlingsansvarlig	*Høgskolen i Oslo og Akershus, ved institusjonens øverste leder*
Daglig ansvarlig	*Sverre Pettersen*
Student	*Magnus Haakens*

Personvernombudet har vurdert prosjektet, og finner at behandlingen av personopplysninger vil være regulert av § 7-27 i personopplysningsforskriften. Personvernombudet tilrår at prosjektet gjennomføres.

Personvernombudets tilråding forutsetter at prosjektet gjennomføres i tråd med opplysningene gitt i **meldeskjemaet, korrespondanse med ombudet, ombudets kommentarer samt personopplysningsloven og helseregisterloven med forskrifter. Behandlingen av personopplysninger kan settes i gang.**

Det gjøres oppmerksom på at det skal gis ny melding dersom behandlingen endres i forhold til de opplysninger som ligger til grunn for personvernombudets vurdering. Endringsmeldinger gis via et eget skjema, http://www.nsd.uib.no/personvern/meldeplikt/skjema.html. Det skal også gis melding etter tre år dersom prosjektet fortsatt pågår. Meldinger skal skje skriftlig til ombudet.

Personvernombudet har lagt ut opplysninger om prosjektet i en offentlig database, http://pvo.nsd.no/prosjekt.

Personvernombudet vil ved prosjektets avslutning, 28.02.2017, rette en henvendelse angående status for behandlingen av personopplysninger.

Vennlig hilsen

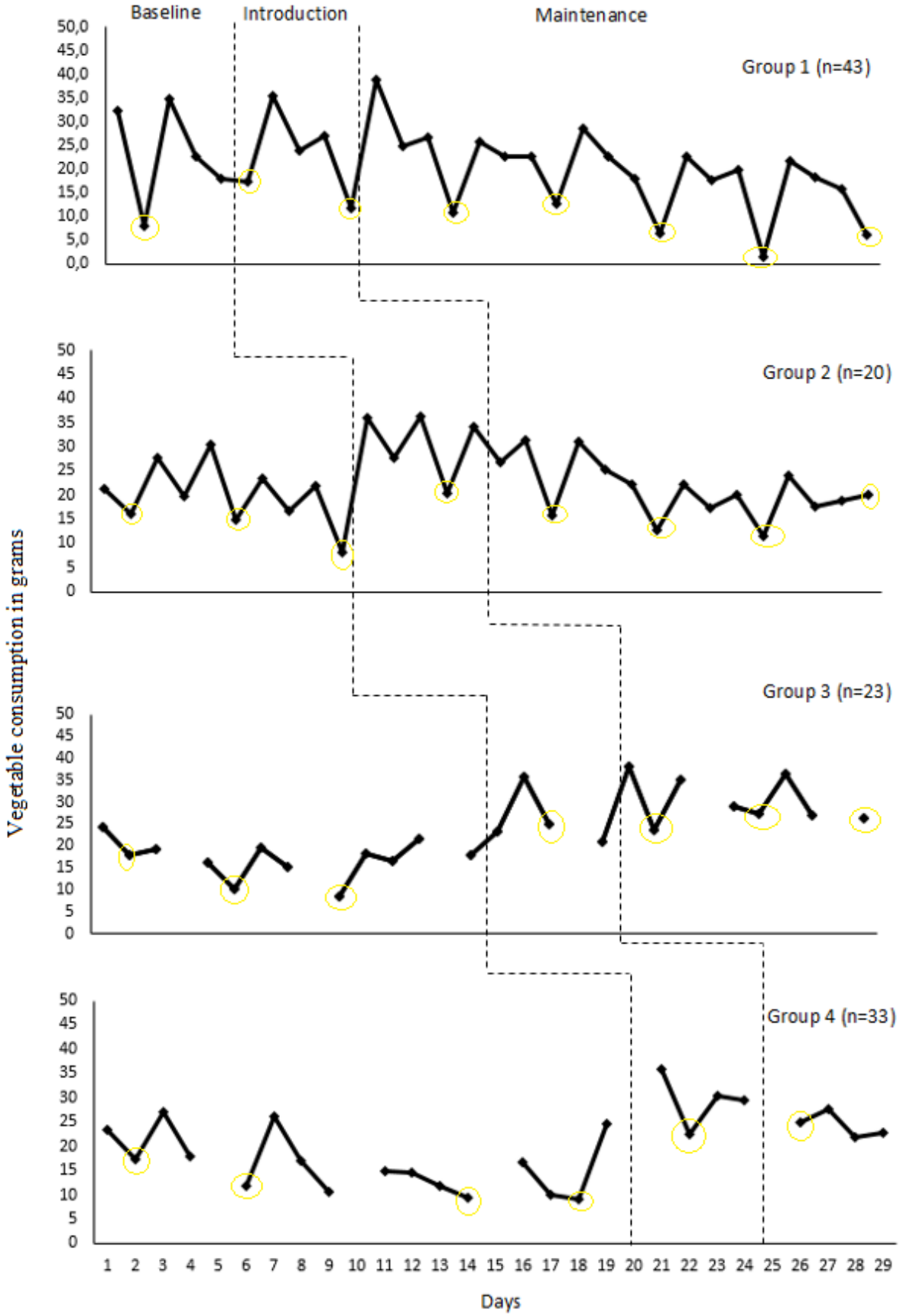
Kjersti Haugstvedt

Siri Tenden Myklebust

Kontaktperson: Siri Tenden Myklebust tlf: 55 58 22 68
Vedlegg: Prosjektvurdering

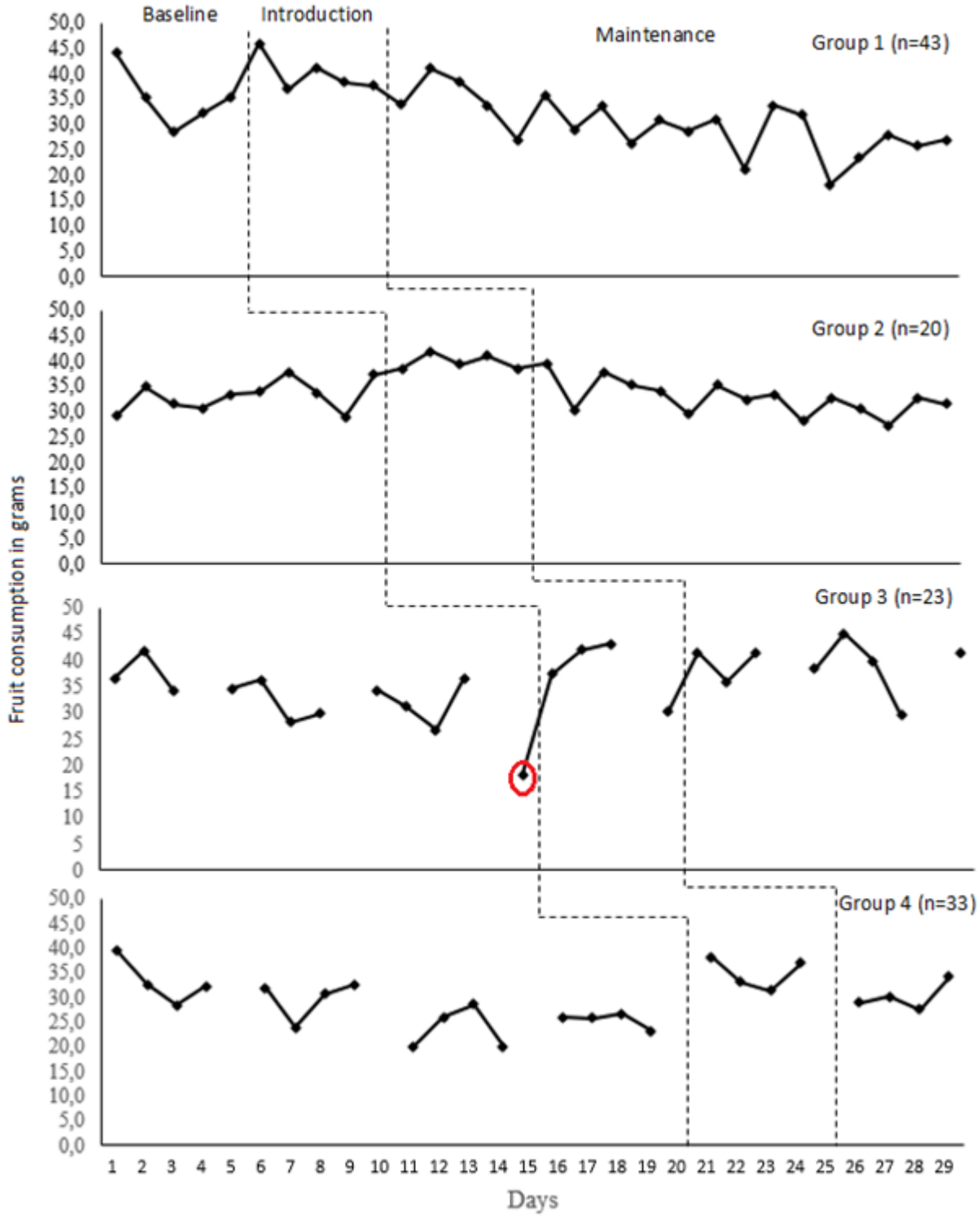
Dokumentet er elektronisk produsert og godkjent ved NSDs rutiner for elektronisk godkjenning.

Appendix 8
Daily average vegetable consumption



Yellow rings indicate days the pupils received rutabaga.

Appendix 9
Daily average fruit consumption



Possible outlier highlited with a red circle.