Evaluering av the Good Behavior Board Game, en variant av the Good Behavior Game

Børge Strømgren
Høgskolen i Oslo og Akershus

og

Dag Gladmann Sørheim
Oslo kommune

Denne studien har testet effekten av Good Behavior Board Game, et tiltak for å skape ro i klasserommet. Good Behavior Board Game er et spill som læreren spiller med klassen. Spillet går ut på at læreren markerer regelbrudd ved å flytte en brikke. Dersom det begås færre enn et gitt antall regelbrudd, tjener elevene tilgang til en hyggelig aktivitet. Regelbrudd er å snakke uten lov, å forlate plassen sin uten lov, og å berøre andre elevers kropp og/eller eiendeler. Tidligere studier har vist at spillet har gode kortsiktige og langsiktige effekter, og lærere og elever gir det gode skussmål, noe som tyder på god sosial validitet. Spillet ble prøvd på en skole der elever og lærere på tre trinn deltok. To av trinnene hadde god effekt av tiltaket, mens hos det tredje fant vi ingen effekt. Mulige årsaker til gode og manglende effekter blir gjennomgått.

Nøkkelord: elevatferd, uro, forstyrrelser, regelbrudd, helklasse skoletiltak, læreratferd, Good Behavior Game, Good Behavior Board Game

---

Elevers uro og forstyrrelse av undervisningen er hyppig forekommende. Opptil 11 % av elevene på fjerde og syvende trinn viser slik atferd i timen i den grad at lærerens undervisning forstyrres, noe som går ut over egen og andres læring (Sørlie & Nordahl, 1998). Situasjonen er ikke mye bedret siden Sørlie og Nordahl sin rapport. To studier gjennomført med 10 års mellomrom—Ogden (1998) og Sørlie og Ogden (2014)—viser stabil gjennomsnittlig prevalens av uro og forstyrrelser i perioden. Også elevundersøkelsen 2012 viser stabilitet over tid med hensyn på elevenes vurdering av variabler som hvor mye tid læreren må bruke på å få ro i klassen, og egen og medelevers uro i timen (Wendelborg, Paulsen, Røe, Valenta & Skaalvik, 2012). I undersøkelsen kan man lese at 30 prosent av elevene svarer at lærerne ofte eller svært ofte må bruke mye tid på å få ro i klassen.

En måte å få ro i klassen på er å satse på de strategiene Hattie (2009) betegner som effektive, for eksempel å motivere elevene til arbeidsinnsats via dialog, forklaring, repetisjon og oppsummering av undervisningen (Nordahl & Hansen, 2012). Klasseledelse, med klare standarder for undervisningen, håndtering av bråk og uro, lærerledelse og struktur, og en positiv lærer-elev relasjon, er svært effektivt (Hattie, 2009; Nordahl & Hansen, 2012). Derfor er det nå fokus på klasseledelse i norsk skole, og utdanningsdi-

Korrespondanse angående denne artikkelen kan adresseres til Børge Strømgren, Høgskolen i Oslo og Akershus. E-post: Borge.Stromgren@hioa.no

Vi ønsker å takke Constanze Nordenstam, Mia Magnusson, Linda Esse og Cecilie Kaaber for hjelp til datainnsamling. Også en stor takk til lærere og elever i aktuelle klasser på angjeldende skole.

rektoratet har en egen nettside med tips og råd til skoleeier og den enkelte lærer (*Klasseledelse*, 2014).

Forskning viser også at hvis lærerens tilbakemelding til elevene er positiv, støttende og konstruktiv er det svært effektiv (Sulla, 2015). Det er en positiv relasjon mellom elevers uro/forstyrelser og lærers bruk av negativ kritikk; når lærer bruker negativ kritikk ofte øker elevers uro/forstyrrelser i tillegg til at de blir mindre fokusert på skolearbeidet. Motsatt, når lærer bruker mye positiv tilbakemelding på faglig innsats og ønsket oppførsel i klasserommet så minker uro/forstyrrelser og akademisk fokus øker. En slik tilbakemelding er essensielt for en positiv elev-lærer relasjon. Til tross for dette er det ikke vanlig med høy forekomst av positive tilbakemeldinger til elever i klasserommet, og det ser ut til å være unaturlig for lærer å gi slike. Forekomsten har også holdt seg stabil over tiår, og lærere blir mer sparsomme med positive tilbakemeldinger og mer rause med negative tilbakemeldinger oppover i klassetrinnene. Forholdet positiv : negativ kan være så mye som 1 : 4, og positive tilbakemeldinger rettes mot akademisk innsats mens negativ kritikk rettes mot uro/forstyrrelser (Sulla).

Sulla (2015) angir grunner til denne situasjonen, at det ser ut til å være unaturlig for lærer å gi positive tilbakemeldinger gitt ro og arbeid med skolearbeid hos elever. Rent bortsett fra at noen lærere kan ha en regel om at positive tilbakemeldinger eller ros ikke burde være nødvendig eller endog kan være skadelig, kan det også forklares med at lærere ikke opplever umiddelbare fordeler ved å gi positive tilbakemeldinger på ønsket elevatferd. Hvis læreren gir ros for at elever sitter rolig og arbeider vil elevene sannsynligvis fortsette med å sitte rolig og arbeide. Dette medfører ingen endringer som kan sees umiddelbart, det vil si at miljøet ikke har blitt endret som følge av rosen. Hvis eleven derimot forstyrrer og læreren gir en korrigerende (negativ) tilbakemelding til eleven, kan det hende eleven stopper å forstyrre—der og da. Læreren opplever en umiddelbar endring i miljøet til det bedre, et eksempel på negativ forsterkning av lærerens atferd. Elevens atferd kan ha blitt positivt forsterket av slik oppmerksomhet, noe som kan bety at forstyrrelser øker på lang sikt—som igjen vil medføre en økning i lærerens korrigerende tilbakemeldinger (Sulla).

Det kan også være at elevers uro/forstyrrelser er opprettholdt av lærerens negative tilbakemeldinger (ved at læreren bruker tid på dette) hvilket betyr en utsettelse eller unngåelse av faglige oppgaver, altså negativ forsterkning av elevens atferd (Hanley, Iwata & McCord, 2003; McIntosh, Horner, Chard, Dickey & Braun, 2008; Sulla, 2015). I tillegg kan lærerens atferd bli negativt forsterket dersom elevens uro/forstyrelse opphører der og da. Slik kan elevens og lærerens atferd opprettholdes av en ”negativ forsterkningsfelle” (Solholm, Askeland, Christiansen & Duckert, 2005).

Det kan altså være lettere og mer ”naturlig” for lærere å konsentrere seg om uro/forstyrrelser i undervisningen og sette i gang tiltak for å redusere slik forekomst, enn det vil være å konsentrere seg om å sette pris på ønsket elevatferd og sette i gang tiltak for å øke denne forekomsten. Utfordringen blir å redusere forekomst av uro/forstyrrelser ved å bruke mindre negative tilbakemeldinger og samtidig øke forekomsten av ros, slik at ro og skolearbeid opprettholdes og/eller økes. Et enkelt tiltak som tar sikte på å håndtere forstyrrelser og uro er Good Behavior Game (GBG), først beskrevet av Wolf og kolleger på slutten av sekstitallet (Barrish, Saunders & Wolf, 1969). GBG er basert på en del atferdsanalytiske prinsipper. De inkluderer gruppekontingenser (Little, Akin-Little & O’Neill, 2015)0, DRO/DRL/DRA (Austin & Bevan, 2011; Conyers et al., 2004; LeGray, Dufrene, Sterling-Turner, Olmi & Bellone, 2010), atferdsavtaler/kontingenskontrakter (Hattie, 2009; Little et al., 2015; Simonsen, Fairbanks, Briesch, Myers & Sugai, 2008), spesifikk/kontingent ros og ”post, teach, review and provide feedback on expecta-

tions" (Heng opp [regler], forklar [reglene], minn om [reglene], og gi feedback [ut fra reglene]) (Simonsen et al., 2008). Videre tegnøkonomi, sosiale forsterkere og samarbeid (Hattie, 2009; Simonsen et al., 2008), og konkrete konsekvenser for regelbrudd (Hattie, 2009). GBG består videre av en del innebygde empirisk støttede kjernepraksiser slik som "mystery motivators", "public posting of feedback of a targeted behavior", "team competition", "timed rewards for inhibition (DRO)", "premack principle", "low emotion or 'private' reprimands" (Embry & Biglan, 2008).

GBG presenteres som et spill for klassen (Kellam, Brown, et al., 2008; Wielkiewicz, 1995), og vanligvis spilles det av og til. I starten spilles det ofte 10 minutters økter, og man har kanskje opptil 3 spill, altså 3 x 10 minutter, per dag opp til 3 dager i uka. På sikt kan man utvide til at spillet varer en hel skoletime eller annen definert periode, opp til 3 dager i uka. Det er altså innebygget at det er økter med spill og økter uten i løpet av skoledagen. Spillet har regler, som vanligvis dreier seg om at elevene skal konkurrere om å vise færrest mulig regelbrudd i løpet av økta . Læreren forklarer hva som er regelbrudd og hva som er ønsket istedenfor. De klassiske regelbruddene i GBG er (1) å snakke uten lov, (2) å forlate plassen sin uten lov, og (3) å berøre andre elevers kropp og/eller eiendeler. De klassiske reglene blir da gjensidig utelukkende atferd; (1) rekk opp hånda når du vil si noe, (2) spør om lov før du forlater plassen din, og (3) hold dine armer og bein for deg selv. Og, siden målatferdene er gjensidig utelukkende, registreres vanligvis bare regelbrudd—det ligger innbakt i prosedyren. Klassen kan deles i ulike lag som konkurrerer om færrest mulig regelbrudd for "sitt" lag eller hele klassen kan konkurrere sammen. Regelbrudd underveis i spillet kommenteres av lærer samtidig som lærer markerer/registrerer regelbrudd på en måte som er synlig for elevene, for eksempel på tavla. Ønsket atferd roses underveis. Når en spilleøkt er ferdig kan man starte en ny med en gang, eller senere på dagen, eller en annen dag. De som har færrest regelbrudd etter en økt og i løpet av uken, eller som ligger under et predefinert nivå av tillatte regelbrudd, tjener opp til belønning[1] for økta eller uken. På sikt utvides tiden for hver økt og antall økter per uke, videre endrer man fra konkrete belønninger som snacks til mer aktivitetsbaserte. Man fader også diskriminative stimuli for spillet, slik at det blir vanskeligere for elevene å vite når spillet er i effekt og ikke.

GBG er ett av få empirisk støttede atferdsanalytiske tiltak (Oliver, Wehby & Reschly, 2011; Simonsen et al., 2008), og det er et universelt tiltak, som innebærer at alle elever i klassen mottar tiltaket samtidig. Tingstrom og kolleger (2006) gjorde en review av 29 enkeltstudier i perioden 1969–2002[2]. Av de 29 studiene de gikk gjennom var 21 "single case design" (SDC), to var ikke oppgitt og fem var randomiserte gruppedesign. Av de 21 SCD var ni rettet mot reduksjon av de klassiske regelbruddene i klassen som ble redusert i alle studiene. En studie hadde "on-task" (jobbe med pålagt oppgave) som målatferd, og GBG var mer effektiv enn individuelle kontingenser. Tre studier hadde økning av skolearbeid som målatferd, noe som var virksomt med GBG. Fire studier hadde en kombinasjon av reduserte regelbrudd og økning av "on-task" som målatferder, noe som også ble oppnådd. De fleste studiene rapporterte om sosial validitet og grad av fornøydhet hos lærere og elever, og vurderingene var gjennomgående positive hos alle involverte.

Siden Tingstrom og kolleger (2006) sin review har vi funnet seks nyere SCD-studier med GBG: de skal kort gjennomgås. Lannie og Curdy (2007) så på endringer i elevenes regelbrudd og fant reduksjoner i disse. I tillegg så de på endringer i lærernes atferd

[1]Vi bruker begrepet belønning siden det vanligvis ikke har vært fortatt forsterkerkartlegging i GBG-studier.

[2]Nolan og kolleger (2014) foretok en review med sikte på internasjonal studier med GBG og bruk av GBG med elever med ulike forutsetninger som fremmedspråklige elever, elever med afrikansk-amerikansk bakgrunn. Nolan og kollegers review er usystematisert og kommenteres ikke videre.

overfor elevene—om positive bemerkninger og ros ville øke og om kjefting og negative utsagn ville minke. De fant ingen økning i positive kommentarer—de lå på null i alle betingelser, og negative kommentarer varierte med elevenes regelbrudd og ikke som en funksjon av gjennomføring av GBG. McCurdy og kolleger (2009) brukte en modifisert GBG kalt LBG (Lunchroom Behavior Game) i kantinen på en barneskole (K-6) med formål om å redusere problematferd, hvilket de oppnådde, og elevene mente at tiltaket var akseptabelt. Tanol og kolleger (2010) sammenlignet to versjoner av GBG—GBG-response cost (GBG-RC) og GBG-forsterkning[3] (GBG-F)—med førsteklassinger (Kindergarten). Målet var økning av klassisk regelfølging og reduksjon av klassiske regelbrudd. GBG-RC innebar at alle lag hadde fire stjerner i starten av spillet, og ved regelbrudd ble én stjerne fjernet samtidig som at regelbruddet ble kommentert av læreren – og de andre lagene fikk ros for å følge reglene. De lagene som hadde én eller flere stjerner igjen etter økta og som nådde ukekriteriet fikk standard GBG-belønning. GBG-F innebar at ingen av lagene hadde noen stjerner i starten av spillet, og når de fulgte reglene fikk de en stjerne og ros, mens lag som brøt reglene ble oversett og ikke fikk stjerne. Også her ble det brukt standard GBG-belønning. Begge versjoner var effektive, men GBG-F var litt mer effektiv og litt bedre likt av lærerne.

Også Donaldson og kolleger (2011) brukte GBG med elever i første klasse, nå med "originalversjonen" der poeng gis for standard regelbrudd med sikte på å redusere disse–noe som skjedde. Elevene likte det godt, 78 % ville fortsette å spille resten av året, og de spilte/lekte GBG på eget initiativ i friminutter. GBG er også brukt på ungdomsskoler, Kleinman og kolleger (2011) gjorde en variant der lagene med færrest standard regelbrudd kunne vinne snack som dagsbelønning og pizza som ukebelønning, mens de som hadde mange regelbrudd kunne miste privilegier. Effekten var god, og elevene mente tiltaket var bra: de lærte mer og lettere samtidig som medelever—og faktisk også de selv—oppførte seg bedre. Til slutt—Flower og kolleger (2014) brukte GBG på en ungdomsskole i en klasse for tilrettelagt matematikkundervisning. Her ble også antall regelbrudd redusert, og lærere og elever likte GBG. Elevene likte spesielt at det var lag som konkurrerte.

Det er også forsket på GBG i et mer langsiktig perspektiv, altså effekten mange år etterpå, men nå med hensyn på ulike former for risikoatferd blant ungdom. Til grunn for denne langtidsforskningen ble det brukt "Randomized Control Design" (RCT), der noen elever fikk GBG i første og andre klasse og andre ikke. De første fem RCT-studiene er omtalt i Tingstrom (2006) sin review, og disse elevene er fulgt opp av en gruppe forskere i Baltimore som i 2008 publiserte en rekke artikler der de viser at når disse GBG-elevene var i 20-årsalderen misbrukte de i mindre grad enn kontrollelevene rusmidler/alkohol/tobakk og hadde lavere forekomst av antisosial atferd og risikabel seksuell atferd sammenlignet med kontrollgruppen (Flay, 2009; Kellam, Reid & Balster, 2008; Kellam et al., 2014). GBG-elevene utførte også færre selvmordsforsøk (Katz et al., 2013; Wilcox et al., 2008; Wyman, 2014). De samme GBG-elevene hadde i ungdomstiden færre emosjonelle problemer og atferdsproblemer enn kontrollelevene, og effekten var størst for de som hadde høyest forekomst av atferdsproblemer som barn (Petras et al., 2008; Poduska et al., 2008).

Senere RCT har i stor grad dreid seg om å undersøke om GBG påvirker utvikling av ulike problemer i ungdomstiden som depresjon, angst og utsatthet for mobbing, noe GBG viser seg å gjøre. Gutter og jenter påvirkes imidlertid ulikt da jenter som har spilt GBG i mindre grad er utsatt for relasjonell mobbing, noe som igjen gir mindre forekomst av angst og depresjon. Gutter som har spilt GBG er mindre utsatthet for fysisk plaging, som igjen gir mindre forekomst av

[3]Her bruker vi begrepet forsterkning fordi forfatterne betegnet betingelsen "GBG-reinforcement".

angst og depresjon (Vuijk, van Lier, Crijnen & Huizink, 2007). Witvliet og kolleger (2009) fant noe av det samme i sin RCT, at elever som har spilt GBG har bedre relasjoner seg imellom i form av aksept fra jevnaldrende (peer acceptance) enn de i kontrollgruppa, noe som gir mindre forekomst av utagerende atferd. Denne funksjonen av å spille GBG, at man påvirker utviklingsbaner med hensyn på problematferd er også funnet i to RCT av Leflot og kolleger (2013; 2010). Spesifikt fant de kortvarige og langvarige bedringer i elevers oppmerksomhet og "on-task"-atferd, samt nedgang i forstyrrelser og aggressiv atferd. De mente at denne effekten kom av spill-elementer som redusert forekomst av kritikk og negative bemerkninger fra lærer. På lengre sikt mente de at effekten kom av bedre aksept blant jevnaldrende for "regelbryteren" som følge av at de spiller sammen, noe som igjen ga reduksjon i aggressiv atferd, mest for de som var uroligst i starten. Dion og kolleger (2011) har også i sin RCT funnet god effekt av GBG på elevers konsentrasjon under lesing, og Spilt og kolleger (2013) fant i sin RCT ut at GBG forhindret atferdsproblemer for elever med emosjonelle problemer og de som var mobbeofre, men ikke for elever fra dysfunksjonelle familier.

De fleste SCD- og RCT-artikler har brukt standardversjonen av GBG eller gjort små endringer. For eksempel har én studie kartlagt forsterkerpreferanser (Kleinman & Saigh, 2011) og noen har brukt "response cost" (RC) (Tanol et al., 2010) ved at man på forhånd delte ut stjerner eller poeng som så ble fjernet ved forekomst av regelbrudd.

En variasjon som er mer omfattende er kalt Good Behavior Board Game (GBBG), som er beskrevet av Cipani (2008, 2010) som en DRL-prosedyre. GBBG er for det første et hel-klasse spill, altså ingen gruppeinndeling. Videre vil en vellykket spilleøkt resultere i en synlig token fremfor en liten belønning der og da. Ved fullt tegnøkonomibrett/full opptjening, vil elevene få tilgang til belønning, for eksempel en hemmelig aktivitet som trekkes fra en boks med ulike belønningslapper i. GBBG kan være lettere for lærere å implementere fordi det ikke er nødvendig å dele inn klassen i ekvivalente grupper og behov for å huske hvem som er i hvilken gruppe. Gjennomføringen vil således virke mindre forstyrende på undervisningen. Elementet med registrering av regelbrudd vil være likt, men elementet med belønningsboks med aktivitetsbelønning kan være mer akseptabelt enn spiselige forsterkere, og det er lett for læreren å sikre at de aktivitetene som er i boksen faktisk lar seg gjennomføre.

Vi kjenner ikke til at GBG eller GBBG er brukt i Norge, så denne studien er et bidrag til en replikasjon og en utprøving av GBBG. Vi ønsket å se på effekten Good Behavior Board Game, og var interessert i å finne ut om spillet kunne skape mer ro i klasserommene i betydningen færre klassiske regelbrudd. Samtidig ønsket vi å undersøke om lærerens atferd ble endret mens spillet pågikk, da i form av flere positive kommentarer og ros til/anerkjennelse av elevene, og færre negative kommentarer/kritikk til elevene.

Metode

Deltakere

Deltakere var (kontakt)lærere og elever på første trinn, fjerde trinn og sjuende trinn på en skole i Oslo. Deltakerne ble valgt etter anbefaling fra ledelsen, og ut fra at de aktuelle (kontakt)lærerne ønsket å få mer ro i undervisningen. Deltakerne og elevenes foresatte ble informert om studiens hensikt og hva deltakelse innebar. Studiens hensikt, å bidra til ro i undervisningssituasjonen og dermed et bedre læringsmiljø, ble vurdert som sosialt valid. Endelig ble tiltaket ansett som sosialt valid av de lærerne som skulle gjennomføre det. Foresatte, skolens ledelse, lærere og elver ble også informert om observasjon i klassen.

Setting og utstyr

Tiltaket ble gjennomført i klassenes vanlige klasserom under ordinær undervisning, eksempelvis norsk og matematikk. Tiltaket ble ikke gjennomført i undervis-

ning utenfor klasserommet, for eksempel kroppsøving.

Utstyr som ble brukt i klasserommet for å gjennomføre spillet var (1) plakater av reglene som skulle overholdes, (2) en abakus (kuleramme) som læreren registrerte og markerte regelbrudd med, (3) et nedtellingsur, (4) store legoklosser som fungerte som tokens, og (5) en belønningsboks med aktivitetslapper i.

Målatferder

Det ble undersøkt om lærernes atferd endret seg som følge av tiltaket og om elevenes atferd endret seg som følge av tiltaket.

Målatferder for elevene. Målatferder var *regelbrudd*, som innbefattet følgende GBG-klassiske regelbrudd: 1) å reise seg fra plassen sin uten lov, 2) å snakke uten å ha blitt gitt ordet, og 3) å ikke holde føtter og hender for seg selv—altså ikke berøre andre elever eller deres pult eller oppgaveark/bøker eller eiendeler med egne hender eller føtter. Hvert slikt regelbrudd ble registrert som en forekomst, og dersom en eller flere av atferdene vedvarte, eller flere elever viste en eller flere av atferdene overlappende, ble det registrert som ny forekomst hvert femte sekund.

Målatferder for lærerne. Målatferder var anerkjennelse av elever, negativ tilbakemelding til elever og (3) positiv tilbakemelding til elever.

Anerkjennelse. Dette innebar at lærer ga oppmerksomhet til elever når de rakk opp hånda eller når de spurte om noe, eller annen form for kommunikasjon mellom lærer og elev. I noen skolesammenhenger brukes begrepet "å se eleven".

Negativ tilbakemelding. Dette innebar at lærer irettesatte elever (Dag, nå må du tie stille) eller klassen (Hysj!), eller brukte negativt ladede ord om elever (Dag, nå tøyser du fælt!) eller klassen (Nå er det veldig bråkete her, dere!).

Positiv tilbakemelding. Dette innebar at lærer ga komplimenter eller ros (Supert, Dag!) til elever eller klassen (Nå jobber dere godt!).

Observasjon og registrering

Målatferder ble observert i de situasjonene som er beskrevet under setting. Alle observasjonsøkter varte i 10 minutter. Hver deltager/klasse ble observert minimum én undervisningstime hver dag hvis det var praktisk mulig, og det ble forsøkt å gjøre tre observasjonsøkter i hver time. Observatørene satt bak i klasserommet for å forstyrre minst mulig. I A-fasen og observasjonsøkter uten spill i B-fasen ble observasjonsøkten startet 10 minutter etter oppstart av timen, eller når klassen hadde kommet til ro. I B-fasen med spill startet observasjonsøkten når lærere annonserte oppstart av spill. Etter at en observasjonsøkt var ferdig, startet en ny etter fem minutters observasjonspause eller ved nytt oppstart av nytt spill. Dersom undervisningen, eller timen, stoppet før en observasjonsøkt var ferdig, ble observasjonen stoppet.

Registreringsinstrument. For å måle forekomst av målatferder brukte observatørene en iPad mini, versjon 7.1.1 (11D201), 30 GB, med observasjonsappen Student/Classroom Observation and Analysis (SCOA) installert. SCOA var stilt inn på å avslutte observasjonen automatisk etter 10 minutter.

Observatørenighet. Det ble gjort samtidige observasjoner for å måle observatørenighet (OE) i 45 % av observasjonsøktene. Da satt observatørene ved siden av hverandre slik at de hadde samme utsyn mot lærer og klassen, men allikevel slikt vendt i forhold til hverandre at de ikke så med-observatørens iPad-skjerm. Observatørene avtalte oppstart av intervallet seg imellom ved å sammen telle ned til start av SCOA-appen. OE ble beregnet med formelen laveste antall delt på høyeste forekomst ganger hundre (Cooper, Heron & Heward, 2007). Det var en generell enighet på 86 %. Tabell 1 viser pålitelighet for de enkelte målatferder i hver klasse, med standardavvik, minimum enighet, maksimum enighet og antall observasjonsøkter der enigheten var under 50 % og over 80 %.

Tabell 1. Observatørenighet i de ulike trinnene for de ulike målatferdene.

Deltakere	Atferdskategori	Snitt	SD	Min	Max	Antall økter med OE < 50 %	Antall økter med OE > 80 %
1A							
Elever	Regelbrudd	84.36	(16.052)	47	100	1 av 22	13 av 22
Lærer	Anerkjennelse	79.86	(24.575)	0	100	1 av 22	14 av 22
Lærer	Negativ tilbakemelding	83.73	(19.134)	33	100	1 av 22	15 av 22
Lærer	Positiv tilbakemelding	87.91	(20.658)	50	100	0 av 22	16 av 22
4B							
Elever	Regelbrudd	87.51	(15.800)	33	100	1 av 47	35 av 47
Lærer	Anerkjennelse	82.12	(21.986)	25	100	8 av 47	32 av 47
Lærer	Negativ tilbakemelding	86.62	(15.487)	50	100	1 av 47	31 av 47
Lærer	Positiv tilbakemelding	85.23	(31.809)	0	100	6 av 47	36 av 47
7A							
Elever	Regelbrudd	85.04	(22.042)	33	100	3 av 23	16 av 23
Lærer	Anerkjennelse	86.00	(17.625)	50	100	1 av 23	14 av 23
Lærer	Negativ tilbakemelding	85.52	(26.604)	0	100	4 av 23	17 av 23
Lærer	Positiv tilbakemelding	88.00	(21.171)	33	100	3 av 23	15 av 23

Prosedyre

Elevene ble etter basislinjeperioden (A-fasen) av klasselæreren fortalt at de skulle spille et spill som heter Good Behavior Board Game. De ble presentert for spillets regler, og hva som ble regnet for å være regelbrudd. Disse reglene ble demonstrert for elevene, og de ble skrevet eller tegnet på plakater som hang foran i klasserommet. Elevene rollespilte også eksempler på regelbrudd samt eksempler på regelfølging.

Deretter fikk elevene vite at spillet ville vare i ti minutter og at tiden ville bli tatt med et nedtellingsur. Dersom det forekom regelbrudd i løpet av spillet ville dette bli markert ved at læreren flytter en kule på en abakus (kuleramme) som var plassert oppå kateteret lett synlig for alle. De fikk vite at dersom de klarte å begå færre enn 10 regelbrudd i løpet av omgangen så ville læreren plassere en stor legokloss på kateteret. Etterpå, eller senere i løpet av timen eller skoledagen, ville et nytt spill starte. Etter ti vellykkede timinutters-perioder, det vil si ti legoklosser på kateteret, vil eleven ha tjent en premie.

Premien var at læreren trakk en lapp opp fra en belønningsboks der det stod beskrevet en aktivitet som barna skulle få gjøre. Noen aktiviteter var mer attraktive enn andre, og det var således knyttet stor spenning til hvilken lapp som ble trukket. Eksempler på aktiviteter var "å spise lunsj utendørs", "tegne i femten minutter", "frilek i gymsalen i tjue minutter", "se på en valgfri musikkvideo på Youtube" eller lignende. Aktivitetene ble valgt av de deltagende lærerne, basert på hva de mente kunne passe som premie for sine elever. Introduksjonen tok ca. 30 min, spillet ble gjennomført i henhold til denne, og det ble spilt hver dag.

Opplæring og veiledning av lærere

Lærerne ble lært opp i bruk av spillet ved informasjon og ved at forfatter to var modell i den angjeldende klasse i noen spilleøkter—disse øktene er ikke med i datagrunnlaget. Videre ble lærerne veiledet noe underveis med hensyn på valg av belønninger mv.

Design og dataanalyse

Vi bruker en multippel basislinjedesign med to serier i B-fasen for hver av målatferdene der hver deltagers data presenteres med fjerdetrinn først, deretter sjuendetrinn og førstetrinn til slutt (Figur 1 og 2). I A-fasen vises basislinjedata uten fylte datapunkter. I B-fasen vises datapunkter på samme måte, altså uten fylte datapunkter, for de observa-

sjonsøktene i tiltaksfasen der GBBG *ikke ble spilt*—dette er en fortsettelse av basislinjebetingelser i B-fasen. I Figur 1 og 2 er dataserien gjennomgående betegnet "Basislinje + GBBG ikke spilt". I B-fasen vises videre datapunkter med sort fyll for de observasjonsøktene der GBBG *ble spilt*. I Figur 1 og 2 er dataserien gjennomgående betegnet "GBBG spilt". Alle datapunktene representerer gjennomsnittlig forekomst i hver timinutters observasjonsøkt i den aktuelle klassen den aktuelle dagen. Designen vil i diskusjonen vurderes ut fra nyere foreslåtte designstandarder (Kratochwill et al., 2013).

I tillegg er data analysert med tanke på effektstørrelse i betydningen prosent ikke-overlappende datapunkter ved hjelp av en Tau-U kalkulator som kan brukes gjennom nettlesere, (http://www.singlecaseresearch.org/calculators). Tau-U tar hensyn til eventuell trend i basislinjedata og man kan korrigere for slik trend i beregningen av overlapp mellom A-fasen og B-fasen dersom trenden går i "feil" retning, altså den retningen man ønsker med tiltaket i B-fasen (Parker, Vannest, Davis & Sauber, 2011). Tau-U gir et tall mellom 0 og 1 som refererer til henholdsvis 0 % og 100 % prosent ikke-overlapp, for eksempel vil en Tau-U på .40 si 40 % ikke-overlapp mellom data i A-fasen og data i B-fasen, og en Tau-U på 65 % eller lavere antyder usikker eller liten forskjell, mellom 66 % og 92 % antyder medium til høy forskjell, og 93 % eller mer antyder svært stor forskjell (Parker et al., 2011; Rakap, 2015). I multiple basislinjedesign kan man beregne Tau-U for hver A-B serie og til slutt for hele settet med AB-serier og slik antyde effekten for hele datasettet i designen (Rakap).

Resultater

Endringer i *Regelbrudd* fra A-fasen til B-fasen er vist i Figur 1, venstre panel. For fjerdetrinn ser en for B-fasen at det ikke er noe bedring i de øktene der det spilles, og forskjellen til basislinjen er liten (Tau-U = .16). For sjuendetrinn ser man en nedgang fra basislinjen og til økter der det spilles[4], og nedgangen er svært stor (Tau-U = .93*[5]). For førstetrinn ser en også en nedgang fra basislinjen og til økter der det spilles, og forskjellen er stor (Tau-U = 1*). Sett under ett viser hele datasettet en usikker effekt med en Tau-U = .60*.

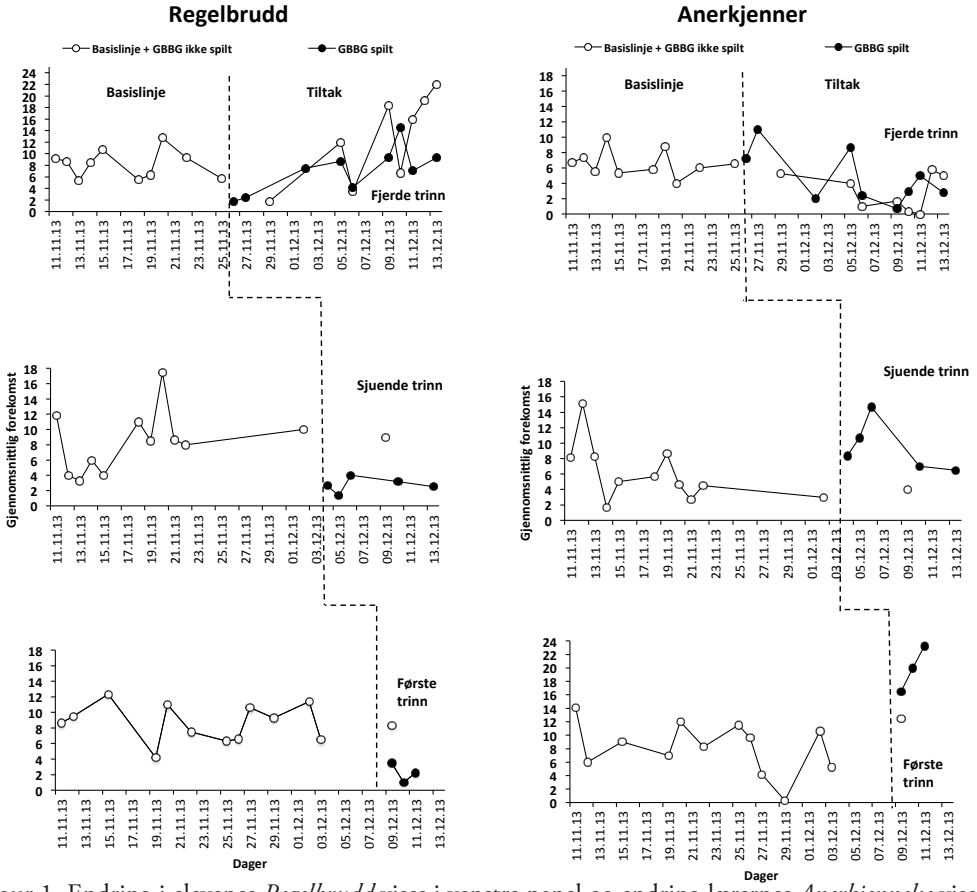
Endringer i *Anerkjennelse* fra A-fasen til B-fasen er vist i Figur 1, høyre panel. For fjerde trinn ser en for B-fasen at anerkjennelse faktisk minker i de øktene der det spilles, selv om nedgangen er liten (Tau-U = .42). For sjuende trinn ser man en økning i økter med spill i B-fasen, og den er også liten (Tau-U = .52). For første trinn ser en at anerkjennelse har økt i økter med spill i B-fasen, og forskjellen er svært stor (Tau-U = 1*). Sett under ett viser hele datasettet en svak effekt med en Tau-U = .20.

Endringer i *Negative kommentarer* fra A-fasen til B-fasen er vist i Figur 2, venstre panel. For fjerdetrinn ser man for B-fasen en nedgang i øktene med spill, og den var middels stor (Tau-U = .68*). For sjuendetrinn ser man for B-fasen en nedgang i øktene med spill, og den var svært stor (Tau-U = 1*). For førstetrinn sees også en nedgang, den var svært stor (Tau-U = 1*). Sett under ett viser hele datasettet en middels til stor effekt med en Tau-U = .85*.

Endringer *Positive kommentarer* fra A-fasen til B-fasen er vist i Figur 2, høyre panel. For fjerdetrinn ser man for det første trend i basislinjen som går i ønsket retning, dataene ble derfor korrigert i Tau-U-beregningen. Med korrigerte basislinjedata er det allikevel i B-fasen en *liten nedgang* i de øktene der det ble spilt (Tau-U = .69*). For sjuendetrinn ser en i B-fasen en økning, som er liten (Tau-U = .40) og for førstetrinn sees også en økning i B-fasen, som er middels stor (Tau-U = .78*). Sett under ett viser hele datasettet ingen effekt med en Tau-U = .01.

[4]Det er ikke gjort sammenligning mellom basislinjen og økter der det ikke spilles for sjuende og første trinn fordi det er for få slike datapunkter i B-fasen.

[5]Når Tau-U-verdien er markert med * er den statistisk signifikant med en $p < .05$.

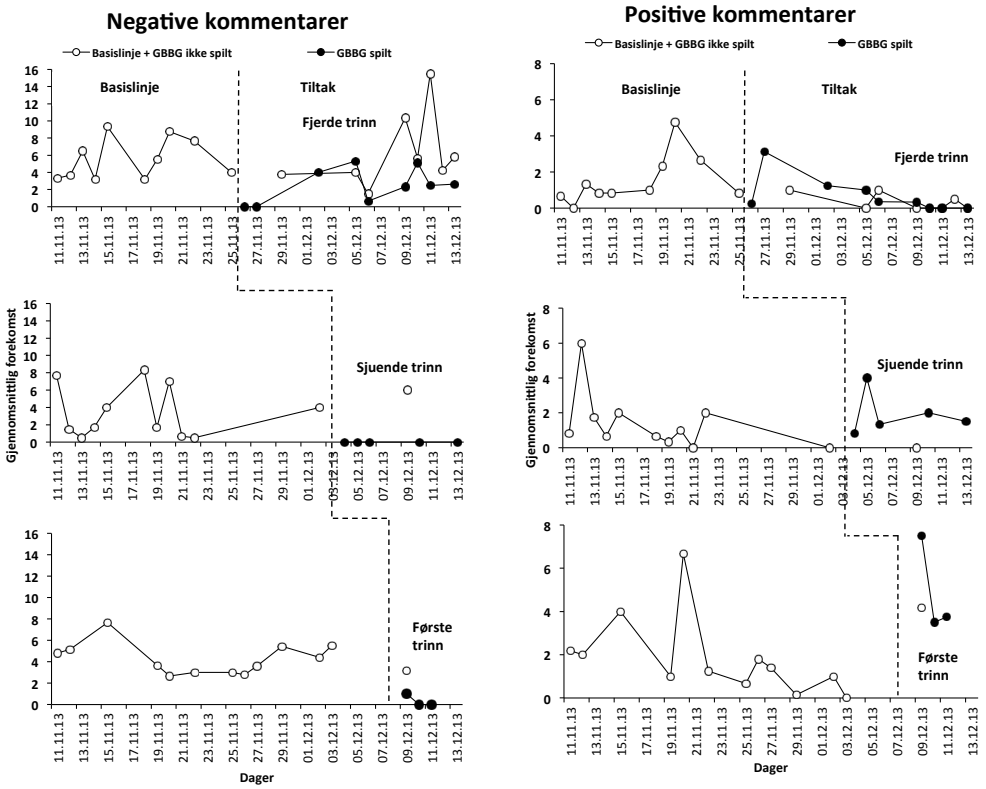


Figur 1. Endring i elevenes *Regelbrudd* vises i venstre panel og endring lærernes *Anerkjennelse* vises i høyre panel. Datapunkter uten fyll representerer basislinjen i A-fasen i tillegg til observasjonsøkter der det **ikke ble spilt** i B-fasen. Datapunkter med sort fyll representerer observasjonsøkter der **det ble spilt** i B-fasen.

Fjerde trinn hadde en lang A-fase, og en tilsvarende lang B-fase med mange datapunkter i alle faser. Over er data fra observasjonsøkter med spill (fylte datapunkter) gjort rede for , men ikke data fra observasjonsøkter uten spill (datapunkter uten fyll). I observasjonsøktene uten spill var det også en økning av *Regelbrudd*, selv om den var liten (Tau-U = .33). Det var også en reduksjon i *Anerkjennelse* i øktene uten spill, og den var middels stor (Tau-U = .86*). For *Negative kommentarer* var det en økning, om enn svak (Tau-U = .13). For *Positive kommentarer* var det en reduksjon som var stor (Tau-U = .91*).

Diskusjon

Vi ønsket å se om Good Behavior Board Game (GBBG) ville være effektivt for å redusere elevenes regelbrudd og samtidig endre noen av lærernes atferder, nærmere bestemt økt anerkjennelse av elevene, mindre negative kommentarer og flere positive kommentarer. Når det gjelder *Regelbrudd* var tiltaket effektivt for sjuende- og førstetrinn, men ikke for fjerdetrinn. Når det gjelder *Anerkjennelse*, økte den for sjuende- og førstetrinn, men ikke for fjerdetrinn. *Negative kommentarer* ble redusert for alle trinn, mens *Positive kommentarer* økte for sjuende- og



Figur 2. Endring i lærernes *Negative kommentarer* vises i venstre panel og endring i lærernes *Positive kommentarer* vises i høyre panel. Datapunkter uten fyll representerer basislinjen i A-fasen i tillegg til observasjonsøkter der det **ikke ble spilt** i B-fasen. Datapunkter med sort fyll representerer observasjonsøkter der **det ble spilt** i B-fasen.

førstetrinn, men avtok for fjerdetrinn.

Resultatene viser at GBBG kan være effektivt, tatt i betraktning de gjennomgående gode resultatene for to av trinnene (sjuende og første), mens de gjennomgående ikke var så gode i ett trinn (fjerde). De gjennomgående dårlige resultatene i fjerdetrinn bidro til å redusere effektstørrelsen for hele datasettet for alle fire målatferdene. Da observatørenighet var nokså lik over trinnene, og også antallet observasjonsøkter, er det sannsynlig at de mindre gode resultatene i fjerde trinn kommer av dårligere gjennomføring av spillet.

Vi har ikke spesifikke data på behandlingstroskap, men observatørene la merke til at spillet i fjerdetrinn ikke ble gjennomført slik som opplæringen hadde lagt opp til. Det kan ikke utelukkes at den aktuelle læreren fikk dårlig opplæring i hvordan tiltaket skulle gjennomføres. Det kan også være at læreren ikke hadde tro på tiltaket, og således gikk halvhjertet inn for det. For fjerdetrinn kan man se at *Negative kommentarer* ikke gikk ned mens spillet ble spilt, og det er et sentralt element i spillet at man istedenfor å påpeke regelbrudd, markerer regelbruddet ved å flytte en kule på abakusen. Det samme ser en for *Positive kommentarer*, ingen økning i spillet slik som forutsatt. Det er også betegnende at data for fjerdetrinn viser at det ikke er noen forskjell mellom observasjonsøkter med spill og observasjonsøkter uten spill.

Hvis man ser på målatferdene *Regelbrudd* og *Negative kommentarer* i B-fasen for fjerdetrinn, ser man at de første datapunktene ser

ut til å gi brukbar effekt, både for observasjonsøkter med og uten spill, men at trenden så går i feil retning. Noe av det samme kan sees på fjerde trinn sine B-fasedata for målatferdene *Anerkjenner* og *Positive kommentarer*, men ikke så uttalt. Slik sett er dette også en indikasjon på svak behandlingstroverdighet, og at den har utviklet seg i uønsket retning underveis. Det kan altså tenkes at læreren for fjerde trinn ikke merket noe bedring i *Regelbrudd* i en tidlig fase av spillet. Etter tredje spill var forekomsten av *Regelbrudd* tilbake på basislinjenivå, noe som kan ha virket som ekstinksjon for å fortsette spillatferden. En ser at bruken av *Negative kommentarer* samtidig øker til basislinjenivå, og at bruken av *Positive kommentarer* fra tredje spill går stabilt nedover. Vi har ikke undersøkt hvorfor læreren ikke fulgte prosedyren, så vi kan ikke si noe om det var mangler ved opplæringen eller ekstinksjonsbetingelser eller andre faktorer som bidro til svak behandlingstroskap. Det er mulig at bedre veiledning underveis og støtte på å "holde ut" kunne ha gitt bedre resultater for fjerde trinn.

Vi kan heller ikke utelukke at resultatet fortsatt hadde vært dårlig selv om tiltaket hadde vært gjennomført på en bedre måte. Det kan for eksempel være at tiltaket i utgangspunktet var uegnet for disse elevene, eller at det ville ta lengre tid før tiltaket virket. Vi har ingen indikasjoner på at det skulle være spesielle forhold ved elevene i fjerdetrinn som skulle tilsi at tiltaket ville ha dårligere effekt her. Imidlertid har det vært en historie med uro i klassen over tid, som også var en av grunnene til at fjerdetrinn var med på forsøket. Det kan være at det er prøvd ulike tiltak i klassen tidligere som har blitt avsluttet uten at de har blitt sluttført, og at klassen sånn sett har lært "å holde ut." På den måten kan det være at man må holde på over tid før en effekt kan påregnes, noe som kan se ut til å være tilfelle, og de siste B-fasedatapunktene for *Regelbrudd* i spilleøktene i kan være en indikasjon på dette.

Det kan også være at resultatet gjennomgående, og for fjerdetrinn spesielt, kunne ha blitt bedre med en mer gjennomtenkt utvelgelse av belønning, eller sågar en preferansekartlegging. Slik preferansekartlegging er lite brukt i GBG, dels fordi det kan være vanskelig å finne en "minste felles multiplum" forsterker for hele klassen, men helst fordi det vanligvis ikke vil være nødvendig. Det er snakk om ordinære elever, såkalt nevrotypiske, i en vanlig klasse på en vanlig skole, og dermed sannsynligvis ikke stort innslag av svært idiosynkratiske forsterkerpreferanser: en kontaktlærer vil i regelen ha en ganske god oversikt over hvilke aktiviteter elevene kunne sette pris på. Selv om Kleinman og Saigh (2011) undersøkte forsterkerpreferanser gjennom uformelle intervjuer var de allikevel avhengige av å bruke stimuli de kalte "kostnadseffektive", altså i tråd med et relativt stramt budsjett, og de brukte konkrete spiselige stimuli slik som snacks og drops etter hvert spill og pizza eller cup-cakes som ukeforsterkere. Dette illustrerer også generelle problemer som kan oppstå med bruk av konkrete stimuli som forsterker i skolen når en hel klasse eller deler av klassen er mottakere: spiselige forsterkere vil være vanskelig med tanke på ulike preferanser og eller diett- og allergirestriksjoner, mens andre konkrete gjenstander vanskelig vil utkonkurrere gjenstander de kan skaffe seg ellers.

Derfor vil aktiviteter, etter Premack's prinsipp, være å foretrekke og det vil være langt enklere å finne en passende aktivitet som belønning[6] for hele klassen. Vi mener dette har fungert bra i denne studien, også fordi den aktivitetene som er opptjent etter 10 tokens (store legoklosser) har vært ukjente på forhånd, da det ble trukket en lapp med aktiviteten fra en belønningsboks slik som i de opprinnelige GBG-studiene (Poduska et al., 2008), og likt som i GBBG der vinnerlaget fikk tilgang til en "hemmelig" aktivitet

[6]Slike aktiviteter mener vi best kan beskrives som belønninger i denne studien, da vi ikke kan påvise en funksjonell relasjon mellom målatferden og produksjon av aktiviteten for alle elevene i klassen, noe som kreves for å bruke begrepet forsterker. Det betyr ikke at en belønning ikke vil være effektiv og at elever ikke kan jobbe for å oppnå den, vi kan bare ikke påvise det eksperimentelt.

(Cipani, 2010), såkalt "mystery motivator" (Embry & Biglan, 2008). Vi mener derfor at den manglende effekten i ett klassetrinn mer er et utslag av selve gjennomføringen av spillet enn av eventuell manglende belønning.

Det er mange kontingenser som kan ha vært effektive for *Regelbrudd* i denne studien. Bruken av kuleramme for å markere regelbrudd kan ha virket som en "offentlig markering av feedback for en offentlig atferd" som man også vet er virksomt (Embry & Biglan, 2008) eller "post, teach, review and provide feedback on expectations" som også er virksomt (Simonsen et al., 2008). Gitt elevenes verbale formuleringer om reglene—at regelbrudd fører til at en kule flyttes på abakusen, og at hvis det blir for mange kuler vil de ikke tjene en legokloss—kan lærerens flytting av kule ha fungert som en betinget positiv straff ved at elevene "tracker" en forverring av situasjonen. Det skal allikevel bemerkes at en eventuell positiv straff som lærerens kuleflytting kan utgjøre sannsynligvis vil være mindre straffende enn en reprimande eller *Negativ kommentar* rettet mot en enkelt elev. Kuleflyttingen kan allikevel være mer effektiv, da som en "low emotion or 'private' reprimand" (Embry og Biglan). I tillegg kan en reprimande fra lærer være både positivt straffende og negativt forsterkende for enkeltelever og slik være mindre treffsikker enn flytting av kule. Det er forskning som tyder på at det er en sammenheng mellom høy forekomst av negative kommentarer fra lærer og mange regelbrudd (Leflot et al., 2010; Sulla, 2015), så kuleflytting eller tilsvarende kan derfor være å foretrekke.

Dersom elevene samarbeider og oppmuntrer hverandres regelfølging i spillet kan det ha fungert som sosial positiv forsterkning, likedan dersom spilleøkten er over uten for mange regelbrudd—elevene kan da ha levert naturlige sosiale forsterkere til hverandre i form av for eksempel utsagn som "Yess, vi klarte det!" og tommelen opp med mere. Elevenes samarbeid om å klare det kan på sikt bedre relasjonen elevene imellom og slik virke inkluderende også for de med mest problemer eller hyppigst regelbrudd i klassen. Dermed får de større anledning til sosial læring fra jevnaldrende og slik lære "normalnivå" av regelbrudd og regelfølging slik som sosialisering vanligvis skjer. Dette kan i følge Witvliet og kolleger (2009) og Leflot og kolleger (2013) forklare en mellomlang effekt av GBG med hensyn på å forhindre utvikling av antisosial atferd, og hvorfor GBG virker best for de som i starten har mest regelbrudd (Spilt et al., 2013). Det kan også tenkes at elevene i løpet av spillet straffet hverandres regelbrudd, for eksempel med utsagn som "Nei, slutt, vi får kule!" eller andre former for sosial misbilligelse, som også er en naturlig del av en sosialiseringsprosess.

Embry og Biglan (2008) bruker betegnelsen "team competition" om slike effektive gruppekontingenser, Simonsen og kolleger (2008) bruker "Class-wide group contingencies". Videre mener Embry og Biglan at elevenes oppmuntring av hverandres regelfølging og misbilligelse av hverandres regelbrudd er dét elementet som sannsynligvis bidrar til en langsiktig preventiv effekt med hensyn på utvikling av problemer og antisosial atferd senere i livet—de lærer hverandre å "holde seg" fra å begå regelbrudd (Selvkontroll) ved at slik "holde seg"-atferd er brakt under umiddelbar negativ forsterkningskontroll og noe mer langsiktig positiv forsterkningskontroll. Derfor bruker Embry (2002) betegnelsen "atferdsvaksine" om GBG. Bruken av legoklosser som tokenbrikke i spillet kan videre ha fungert som betinget forsterker i og med at opptjening for det første signaliserer en vellykket spilleøkt for klassen, og videre at 10 brikker var assosiert med adgang til en overraskelsesaktivitet, en såkalt "mystery motivator" (Embry & Biglan, 2008). Bruken av legoklosser kan også ha fungert som *tracks* som signaliserte en forbedring i situasjonen gitt at elevene hadde lært og kunne beskrive regelen "10 legoklosser betyr overraskelsesaktivitet", ved at læreren hadde beskrevet reglen OG fulgt den.

Det kan også være flere kontingenser for lærernes spillatferd, og i spillatferd inngår også økning i *Anerkjennelse* og *Positive kommentarer* og reduksjon i *Negative kommentarer*. Vi skal peke på noen "naturlig forekommende" kontingenser som kan virke mens lærer spiller. En reduksjon i *Regelbrudd* kan virke som en positiv forsterker for fortsatt spilling i seg selv. Atferden å flytte en kule på abakusen (eller tilsvarende markering) gitt et regelbrudd vil sannsynligvis ikke naturlig forsterkes i starten, det er her snakk om å "utvide operanten" eller å endre fra *Negative kommentarer*. Gitt en læringshistorie der *Negative kommentarer* har redusert antall *Regelbrudd* – om så kortvarig, vil *Negative kommentarer* være negativt forsterket atferd. En overgang til "kuleflytting" vil da i starten sannsynligvis være opprettholdt av spillets regler heller enn av endringer i elevenes *Regelbrudd*. Over repetisjonene vil en reduksjon i *Regelbrudd* kunne opprettholde "kuleflytting" på samme måte som *Negative kommentarer* ble opprettholdt, gjennom negativ forsterkning. En slik utvikling ser man i sjuende- og førstetrinn sine data for *Regelbrudd*, og for disse lærernes data for *Negative kommentarer*.

Gitt at *Regelbrudd* går ned og undervisningen "flyter" bedre kan det bli lettere for læreren å legge merke til de som følger reglene også; de "rolige" elevene, eller regelfølging, kan her fungere som diskrimativ stimulus for å gi ros og lignende. Det kan forklare økningen i *Anerkjennelse*. Vi ser en slik mulig sammenheng for sjuende og første trinn, men ikke for fjerde—her gikk ikke *Regelbrudd* ned og de overskygget muligens elevenes regelfølging som S^D. I tillegg kan naturlige sosiale forsterkere opprettholde *Anerkjennelse*, man ser effekten hos den eleven man anerkjenner, eksempelvis smil og en hyggelig tone med mere. Den samme virkningsmekanismen kan gjelde for *Positive kommentarer*, selv om vi ser liten endring på sjuendetrinn, mye endring på førstetrinn, og reduksjon på fjerdetrinn. *Positive kommentarer* er som før nevnt ikke "naturlig" (Sulla, 2015), og det blir mer unaturlig jo høyere opp i klassetrinnene man kommer. Dataene for *Positive kommentarer* i sjuende og første trinn kan derfor dels gjenspeile en naturlig forskjell mellom trinn, og dels gjenspeile regelfølging i GBBG der *Positive kommentarer* er en del av spillet.

Mens RCT-studiene har forklart effekten av GBG rundt bedringer i elevers relasjoner som følge av at de spiller sammen på et lag—gruppekontingenser, har man i SCD-studiene lagt mer vekt på prosedyremessige forklaringsmodeller som DRO/RC/DRL/DRA og atferdsavtaler/kontingenskontrakter. Embry (2002) har som nevnt en forklaring som ligger mellom disse da han fremholder at det er treningen på å "holde" seg, eller inhibisjon og selvkontroll, som er den virksomme langsiktige komponenten. Videre mener Embry at den kortsiktige komponenten ikke bare kan tilskrives enklere prosedyrer som for eksempel DRO alene, fordi GBG består av en mengde effektive kjernepraksiser som presenteres samtidig.

Selv om GBG originalt er utformet som en DRO/DRL-prosedyre (Barrish et al., 1969), er kjernepraksiser som "mystery motivators", "public posting of feedback of a targeted behavior", "team competition", "timed rewards for inhibition (DRO)", "Premack's principle", "low emotion or 'private' reprimands", innebygd (se for eksempel Embry og Biglan, 2008), og de modifiserer muligens innslaget av DRO/DRL-prosess i tiltaket, selv om prosedyren fremdeles har en DRO/DRL-logikk. Videre kan en direkte muntlig, og praktisk rollespill-basert, gjennomgang med elevene om hvilke atferder man søker å redusere—og hvilke som bør øke—ytterligere redusere innslaget av DRO-prosess. Tiltaket kan da bli mer likt DRA fordi de vet hvilken atferd de må vise for å oppnå ønsket konsekvens (Fredheim & Finstad, 2006), altså nærmere en positiv forsterkningsprosedyre enn en negativ straffeprosedyre—også fordi det ligger innbakt at lærer skal levere flere *Positive* enn *Negative kommentarer*.

Vi synes det er vanskelig å si noe sikkert om GBG/GBBG er en negativ straffeprosedyre som ved standard DRO/DRL eller ikke. Sannsynligvis er det mange samtidige kontingenser i effekt underveis, slik som i livet ellers. Utenfor laboratoriet er det aldri kun én kontingens i effekt. Vi mener en styrke ved spillet er at det mimer det naturlige livet. Det er ikke naturlig å fortsette å forsterke ønsket atferd inn i evigheten, på ett tidspunkt sluttet vi å få ros fra vår mor for å knytte skolissene våre, og det gikk etter hvert over til påpakning dersom vi ikke hadde gjort det. Slike kontingensskifter skjer naturlig i oppveksten og i mellommenneskelige relasjoner, så også i skolen og i det å være skoleelev.

I tillegg er det ikke slik at all problematferd vil opphøre dersom man kun vektlegger å rose uforenlig atferd, og om det skulle skje vil det ta lang tid. Kanskje for lang tid for de elevene med atferdsproblemer som kunne risikert eksklusjon fra jevnaldrende underveis. Så selv om GBG/GBBG best kan betegnes som en negativ straffeprossedyre mener vi at det ikke utgjør noe problem i den formen den her forekommer. Riktig bruk vil redusere negative kommentarer fra lærer og øke positive, noe som kan utligne de negative straffeelementene som kuleflytting og eventuell manglende opptjening av token kan utgjøre. I tillegg er det ikke nødvendigvis et sentralt poeng å fjerne alle sosiale stimuli som kan minne om negativ straff, det viktige er at forholdet positive : negative kommentarer er større en forholdet negative : positive. Forskningen viser at det i regelen er mer negative kommentarer i skolen enn positive, opp til fire negative på en positiv (Sulla, 2015), og GBBG er ett bidrag for å snu dette i mer ønsket retning.

Vi kan heller ikke se at spillet som helhet har fungert som en negativ straffeprosedyre i denne studien. Atferd som ofte assosieres med straff, slik som økning i problematferd og motkontroll, har vært helt fraværende, elevene var i to av trinnene svært entusiastiske og ville spille mer enn lærer ga anledning til. Slik sett ser det ut til at den sosiale validiteten har vært like god som i tidligere studier. Igjen, den manglende reduksjonen i *Regelbrudd* man så på det ene trinnet mener vi er et resultat av at spillet ikke ble korrekt utført og ikke av en eventuell straffeindusert effekt—økningen i *Regelbrudd* kom i de observasjonsøktene der det ikke ble spilt. Dette kan tyde på at spillet relativt sett har hatt en viss effekt for fjerdetrinn. Læreren for dette trinnet sluttet også å spille, mens de to andre lærerne sa seg så fornøyd med spillet at de fortsatte også etter studien, noe vi mener også vitner om god sosial validitet.

Noen svakheter og styrker skal kommenteres. Ut fra nyere designstandarder (Kratochwill et al., 2013) kan en si at designen holder en "standard med reservasjoner". Blant annet er det ikke fem datapunkter i alle fasene. Det var for eksempel ikke nok tid til å få til dette for førstetrinn, der klarte vi fire datapunkter (tre med spillet og ett uten spillet). For datasettet *Anerkjenner* er det også en dataserie med ikke-effekt, og det er en dataserie med liten effekt for datasettet *Positive kommentarer*. Det er allikevel nok til å vurdere evidensstyrken i studien, noe vi mener er en styrke ved studien.

Vi har valgt å vurdere evidensstyrken ut fra trend da vi mener det best gjenspeiler hvordan en kan forvente en atferdsendring på et helt trinn, og hos lærere, når det er på bakgrunn av korte perioder med spill. Det ble brukt kuleramme og nedtellingsur under spillet og ikke ellers. Det ble derfor lett for elevene å diskriminere mellom når spillet ble spilt og ikke. Det kan ta noe tid før nye kontingenser virker, og vi mener derfor at en vurdering av trend er mer riktig enn for eksempel rask endring i nivå. For de fleste basislinjer, unntatt *Regelbrudd* og *Negative kommentarer* på førstetrinn, ser en at basislinjetrenden går i motsatt retning av ønsket endring av tiltaket.

For datasettet *Regelbrudd* ser en for sjuendetrinn og førstetrinn at basislinjetrenden fortsetter inn i tiltaksfasen når det ikke spilles (åpne hvite datapunkter i B-fasen), men det

er en nedgang når det spilles (lukkede sorte datapunkter). På fjerdetrinn ser man ikke en klar forskjell før de tre siste datapunkter i B-fasen når en tar i betraktning kontrasten mellom datapunkter med (lukkede sorte) og uten spill (åpne hvite). Dette inntrykket fra den visuelle analysen støttes av Tau-U-analysene. Det samme hovedbildet gjør seg gjeldende for de øvrige datasettene og tilhørende Tau-U-analyser. Vi vil derfor hevde at designen har rimelig god evidensstyrke for to av datasettene, *Regelbrudd* og *Negative kommentarer*, men ikke for de andre. Dette mener vi er en svakhet ved studien som helhet. Vi har allikevel valgt å også ta med datasett med ikke-effekt eller usikker effekt.

Senere studier bør systematisk ta data på behandlingsintegritet og legge til rette for raskt tilgjengelig opplæring og veiledning når data viser behov for det. Slik kan man få en bedre test av GBBG. Andre viktige implementeringsfaktorer kan være å planlegge bedre når det skal spilles og slik få anledning til å samle data i tiltaksfasen både når det spilles og når det ikke spilles. Det er imidlertid en utfordring å planlegge for dette i skolen, da det er mye som skjer og fordi planer ofte må endres. Forskning ute i vanlig skole vil naturlig nok påvirkes av at det er vanskeligere å kontrollere betingelsene.

En måte å løse dette på er å legge opp til en lengre observasjonsperiode; for eksempel kan 6 uker med observasjon være noe kort tid. Vi skulle gjerne sett hvilket tiltak som var mest effektivt av GBG (med gruppeinndeling av klassen) og GBBG (helklasse). Vi vil derfor anbefale at det gjennomføres en studie der man sammenligner effektene av de to tiltakene. Vi anbefaler også bruk av nettbrett til å samle data, det kan gi sikrere data og også være med å redusere observatøreffekten, da det blir mindre tydelig for omgivelsene når det faktisk registreres i løpet av timen enn hvis man bruker penn og papir.

Oppsummert vi vil hevde at GBBG er et tiltak som kan være effektivt for å redusere uro og forstyrrelser (*Regelbrudd*) i vanlige klasser med vanlige elever. Det er relativt lett å gjennomføre, lærere kan gjøre dette samtidig med vanlig undervisning, og underveis i spillet trenger de kun å markere regelbrudd og å sikre ros ved elevers regelfølging. Det trengs lite og rimelig utstyr, og aktiviteter kan brukes som belønning (back-up-forsterkere). Bruk av spillet kan gi mer tid til skolerettet arbeid og effektiv undervisning, og mindre tid brukt til å skape ro og ordne opp i konflikter elevene imellom. Spillet er trygt å gjennomføre og lærere og elever setter pris på det.

Referanser

Austin, J. L. & Bevan, D. (2011). Using differential reinforcement of low rates to reduce children's requests for teacher attention. *Journal of Applied Behavior Analysis, 44*, 451–461. doi:10.1901/jaba.2011.44-451

Barrish, H. H., Saunders, M. & Wolf, M. M. (1969). Good Behavior Game: Effects of individual contingencies for group consequences on disruptive behavior in a classroom. *Journal of Applied Behavior Analysis, 2*, 119–124. doi:10.1901/jaba.1969.2-119

Cipani, E. (2008). *Classroom management for all teachers: 12 plans for evidence-based practice* (3. utg.). Columbus, OH: Prentice Hall.

Cipani, E. (2010). The class-wide Good Behavior Board Game. Hentet fra http://eric.ed.gov/?id=ED512078

Conyers, C., Miltenberger, R., Maki, A., Barenz, R., Jurgens, M., Sailer, A., . . . Kopp, B. (2004). A comparison of response cost and differential reinforcement of other behavior to reduce disruptive behavior in a preschool classroom. *Journal of Applied Behavior Analysis, 37*, 411–415. doi:10.1901/jaba.2004.37-411

Cooper, J. O., Heron, T. E. & Heward, W. L. (2007). *Applied behavior analysis*. Upper Saddle River, NJ: Pearson Education, Inc.

Dion, E., Roux, C., Landry, D., Fuchs, D., Wehby, J. & Dupéré, V. (2011). Improving attention and preventing reading

difficulties among low-income first-graders: A randomized study. *Prevention Science, 12*, 70–79. doi:10.1007/s11121-010-0182-5

Donaldson, J. M., Vollmer, T. R., Krous, T., Downs, S. & Berard, K. P. (2011). An evaluation of the Good Behavior Game in kindergarten classrooms. *Journal of Applied Behavior Analysis, 44*, 605–609. doi:10.1901/jaba.2011.44-605

Embry, D. D. (2002). The Good Behavior Game: A best practice candidate as a universal behavioral vaccine. *Clinical Child and Family Psychology Review, 5*, 273–297. doi:10.1023/A:1020977107086

Embry, D. D. & Biglan, A. (2008). Evidence-based kernels: Fundamental units of behavioral influence. *Clinical Child and Family Psychology Review, 11*, 75–113. doi:10.1007/s10567-008-0036-x

Flay, B. (2009). School-based smoking prevention programs with the promise of long-term effects. *Tobacco Induced Diseases, 5*. doi:10.1186/1617-9625-5-6

Flower, A., McKenna, J., Muething, C. S., Bryant, D. P. & Bryant, B. R. (2014). Effects of the Good Behavior Game on classwide off-task behavior in a high school basic algebra resource classroom. *Behavior Modification, 38*, 45–68. doi:10.1177/0145445513507574

Fredheim, T. & Finstad, J. (2006). Negativ straff i form av response cost og DRO i behandling av problematferd. Teori og praksis. *Norsk Tidsskrift for Atferdsanalyse, 33*, 71–81. Hentet fra http://nta.atferd.no/journalissue.aspx?IdDocument=204

Hanley, G. P., Iwata, B. A. & McCord, B. E. (2003). Functional analysis of problem behavior: A review. *Journal of Applied Behavior Analysis, 36*(2), 147–185. doi:10.1901/jaba.2003.36-147

Hattie, J. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achivement*. London: Routledge.

Katz, C., Bolton, S.-L., Katz, L. Y., Isaak, C., Tilston-Jones, T. & Sareen, J. (2013). A systematic review of school-based suicide prevention programs. *Depression and anxiety, 30*, 1030–1045. doi:10.1002/da.22114

Kellam, S. G., Brown, C. H., Poduska, J. M., Ialongo, N. S., Wang, W., Toyinbo, P., . . . Wilcox, H. C. (2008). Effects of a universal classroom behavior management program in first and second grades on young adult behavioral, psychiatric, and social outcomes. *Drug and Alcohol Dependence, 95, Supplement 1*, S5–S28. doi:http://dx.doi.org/10.1016/j.drugalcdep.2008.01.004

Kellam, S. G., Reid, J. & Balster, R. L. (2008). Effects of a universal classroom behavior program in first and second grades on young adult problem outcomes. *Drug and Alcohol Dependence, 95, Supplement 1*, S1–S4. doi:http://dx.doi.org/10.1016/j.drugalcdep.2008.01.006

Kellam, S. G., Wang, W., Mackenzie, A. C. L., Brown, C. H., Ompad, D. C., Or, F., . . . Windham, A. (2014). The impact of the Good Behavior Game, a universal classroom-based preventive intervention in first and second grades, on high-risk sexual behaviors and drug abuse and dependence disorders into young adulthood. *Prevention Science, 15*, 6–18. doi:10.1007/s11121-012-0296-z

Klasseledelse. (2014). Hentet 28.05.2014 fra http://www.udir.no/Laringsmiljo/Bedre-laringsmiljo/Klasseledelse/

Kleinman, K. E. & Saigh, P. A. (2011). The effects of the Good Behavior Game on the conduct of regular education New York City high school students. *Behavior Modification, 35*, 95–105. doi:10.1177/0145445510392213

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M. & Shadish, W. R. (2013). Single-case intervention research design standards. *Remedial and Special Education, 34*, 26–38. doi:10.1177/0741932512452794

Lannie, A. L. & McCurdy, B. L. (2007). Preventing disruptive behavior in the urban classroom: Effects of the Good

Behavior Game on student and teacher behavior. *Education and Treatment of Children (West Virginia University Press), 30*, 85–98. doi:10.1353/etc.2007.0002

Leflot, G., van Lier, P. A. C., Onghena, P. & Colpin, H. (2013). The role of children's on-task behavior in the prevention of aggressive behavior development and peer rejection: A randomized controlled study of the Good Behavior Game in belgian elementary classrooms. *Journal of School Psychology, 51*, 187–199. doi: 10.1016/j.jsp.2012.12.006

Leflot, G., van Lier, P. C., Onghena, P. & Colpin, H. (2010). The role of teacher behavior management in the development of disruptive behaviors: An intervention study with the Good Behavior Game. *Journal of Abnormal Child Psychology, 38*, 869–882. doi:10.1007/s10802-010-9411-4

LeGray, M., Dufrene, B., Sterling-Turner, H., Olmi, D. J. & Bellone, K. (2010). A comparison of function-based differential reinforcement interventions for children engaging in disruptive classroom behavior. *Journal of Behavioral Education, 19*, 185–204. doi:10.1007/s10864-010-9109-2

Little, S. G., Akin-Little, A. & O'Neill, K. (2015). Group contingency interventions with children—1980-2010: A Meta-Analysis. *Behavior Modification, 39*, 322–341. doi:10.1177/0145445514554393

McCurdy, B. L., Lannie, A. L. & Barnabas, E. (2009). Reducing disruptive behavior in an urban school cafeteria: An extension of the Good Behavior Game. *Journal of School Psychology, 47*, 39–54. doi:10.1016/j.jsp.2008.09.003

McIntosh, K., Horner, R. H., Chard, D. J., Dickey, C. R. & Braun, D. H. (2008). Reading skills and function of problem behavior in typical school settings. *The Journal of Special Education, 42*, 131–147. doi:10.1177/0022466907313253

Nolan, J. D., Houlihan, D., Wanzek, M. & Jenson, W. R. (2014). The Good Behavior Game: A classroom-behavior intervention effective across cultures. *School Psychology International, 35*, 191–205. doi:10.1177/0143034312471473

Nordahl, T. & Hansen, O. (2012). *Klasseledelse*. Oslo: Gylendal Akademisk.

Ogden, T. (1998). *Elevatferd og læringsmiljø. Læreres erfaringer med og syn på elevatferd og læringsmiljø i grunnskolen* (Rapport '98). Oslo: Kirke-, utdannings- og forskningsdepartementet.

Oliver, R. M., Wehby, J. H. & Reschly, D. J. (2011). *Teacher classroom management practices: Effects on disruptive or aggressive student behavior: A systematic Review* (Campbell Systematic Reviews 2011.4). doi:10.4073/csr.2011.4

Parker, R. I., Vannest, K. J., Davis, J. L. & Sauber, S. B. (2011). Combining nonoverlap and trend for single-case research: Tau-U. *Behavior Therapy, 42*, 284–299. doi:10.1016/j.beth.2010.08.006

Petras, H., Kellam, S. G., Brown, C. H., Muthén, B. O., Ialongo, N. S. & Poduska, J. M. (2008). Developmental epidemiological courses leading to antisocial personality disorder and violent and criminal behavior: Effects by young adulthood of a universal preventive intervention in first- and second-grade classrooms. *Drug and Alcohol Dependence, 95, Supplement 1*, S45–S59. doi:http://dx.doi.org/10.1016/j.drugalcdep.2007.10.015

Poduska, J. M., Kellam, S. G., Wang, W., Brown, C. H., Ialongo, N. S. & Toyinbo, P. (2008). Impact of the Good Behavior Game, a universal classroom-based behavior intervention, on young adult service use for problems with emotions, behavior, or drugs or alcohol. *Drug and Alcohol Dependence, 95S*, S29–S44. doi:10.1016/j.drugalcdep.2007.10.009

Rakap, S. (2015). Effect sizes as result interpretation aids in single-subject experimental research: Description and application of four nonoverlap methods. *British Journal of Special Education, 42*, 11–33. doi:10.1111/1467-8578.12091

Simonsen, B., Fairbanks, S., Briesch, A., Myers, D. & Sugai, G. (2008). Evidence-based practices in classroom management: Considerations for research to practice. *Education & Treatment of Children, 31*(3), 351–380. Hentet fra http://search.proquest.com/docview/202675482?accountid=26439

http://openurl.bibsys.no/openurl?issn=07488491&aulast=Simonsen&aufirst=Brandi&title=Education+%26+Treatment+of+Children&atitle=Evidence-based+Practices+in+Classroom+Management%3A+Considerations+for+Research+to+Practice1&volume=31&issue=3&date=2008&pages=351&isbn=

Solholm, R., Askeland, E., Christiansen, T. & Duckert, M. (2005). Parent Management Training – Oregon-modellen. *Tidsskrift for Norsk Psykologforening, 42*, 587–597. Hentet fra http://www.psykologtidsskriftet.no/?seks_id=297106&a=2

Spilt, J., Koot, J. & van Lier, P. C. (2013). For whom does it work? Subgroup differences in the effects of a school-based universal prevention program. *Prevention Science, 14*, 479–488. doi:10.1007/s11121-012-0329-7

Sulla, F. (2015). *The effect of teacher approval on students social and academic behavior* (Unpublished doctoral dissertation). University of Parma, Parma.

Sørlie, M. A. & Nordahl, T. (1998). *Problematferd i skolen. Hovedfunn, forklaringer og pedagogiske implikasjoner. Hovedrapport fra forskningsprosjektet "Skole og samspillsvansker"*. Oslo: Norsk institutt for forskning om oppvekst, velferd og aldring.

Sørlie, M. A. & Ogden, T. (2014). Mindre problematferd i grunnskolen? - Lærervurderinger i et 10-års perspektiv. *Norsk pedagogisk tidsskrift, 98 E*, 190–202. Hentet fra http://www.idunn.no/ts/npt/2014/03/mindre_problematferd_igrunnskolen_-_laerervurderinger_i_et

Tanol, G., Johnson, L., McComas, J. & Cote, E. (2010). Responding to rule violations or rule following: A comparison of two versions of the Good Behavior Game with kindergarten students. *Journal of School Psychology, 48*, 337–355. doi:10.1016/j.jsp.2010.06.001

Tingstrom, D. H., Sterling-Turner, H. E. & Wilczynski, S. M. (2006). The Good Behavior Game: 1969–2002. *Behavior Modification, 30*, 225–253. doi:10.1177/0145445503261165

Vuijk, P., van Lier, P. A. C., Crijnen, A. A. M. & Huizink, A. C. (2007). Testing sex-specific pathways from peer victimization to anxiety and depression in early adolescents through a randomized intervention trial. *Journal of Affective Disorders, 100*, 221–226. doi:http://dx.doi.org/10.1016/j.jad.2006.11.003

Wendelborg, C., Paulsen, V., Røe, M., Valenta, M. & Skaalvik, E. (2012). *Elevundersøkelsen 2012*. Trondheim: NTNU Samfunnsforskning AS.

Wielkiewicz, R. M. (1995). *Behavior management in the schools: Principles and procedures* (2. utg.). Needham Heights, MA Allyn & Bacon.

Wilcox, H. C., Kellam, S. G., Brown, C. H., Poduska, J. M., Ialongo, N. S., Wang, W. & Anthony, J. C. (2008). The impact of two universal randomized first- and second-grade classroom interventions on young adult suicide ideation and attempts. *Drug and Alcohol Dependence, 95, Supplement 1*, S60–S73. doi:http://dx.doi.org/10.1016/j.drugalcdep.2008.01.005

Witvliet, M., van Lier, P. A. C., Cuijpers, P. & Koot, H. M. (2009). Testing links between childhood positive peer relations and externalizing outcomes through a randomized controlled intervention study. *Journal of Consulting and Clinical Psychology, 77*, 905–915. doi:10.1037/a0014597

Wyman, P. A. (2014). Developmental approach to prevent adolescent suicides: Research pathways to effective upstream preventive interventions. *American Journal of Preventive Medicine, 47*, S251–S256. doi:http://dx.doi.org/10.1016/j.amepre.2014.05.039

Evaluation of the Good Behavior Board Game: A variation of the Good Behavior Game

Børge Strømgren, Oslo and Akershus University College
and
Dag Gladmann Sørheim, City of Oslo

The current study tested the effect of the Good Behavior Board Game, an intervention aimed towards creating order in the classroom. The teacher played the Good Behavior Board Game with the class, and during the play the teacher marked rule violations by moving a bead on an abacus. If fewer than an agreed upon number of rule violations were committed, the students earned an activity. Rule violations included talking out aloud without permission, leaving seat without permission, and touching other student's body and/or belongings. Earlier studies have shown positive short- and long-term effects on student behavior, and teachers and students have evaluated it positively, which indicates good social validity. In the current study three teachers played the game with their classes. Two of the classes showed a good effect of the game, the third class showed no effect. Possible reasons for the god and lacking effects are discussed.

Keywords: student behavior, disruptions, rule violations, whole class school intervention, teacher behavior, Good Behavior Game, Good Behavior Board Game