

Högskolepedagogiska texter, Enheten för pedagogik och interaktivt lärande (PIL)
Göteborgs universitet, maj 2016

Bedömning och betygsättning av experimentella grupparbeten

Ann Larsson

Institutionen för Marina vetenskaper

Göteborgs universitet

Abstract

Rättvis betygsättning av grupparbeten är en utmaning för oss lärare. Inom mitt ämnesområde, marin ekologi, är det vanligt med grupparbeten och jag undervisar på en 15 hp kurs där ett 6 veckor långt experimentellt grupparbete ingår. Grupparbetet är obligatoriskt men har inte tidigare specifikt betygsatts och jag undersöker här vilka möjligheter det finns att betygsätta grupparbeten med speciellt fokus på inbördes bedömning (peer evaluation) vilket innebär att studenterna utvärderar varandra inom gruppen vilket ger underlag till individuell betygsättning. Information om olika modeller för inbördes bedömning samlades in genom att studera befintlig litteratur där metoder för poängsättning av varandra i gruppen, hur resultatet av inbördes bedömning implementeras i totala betyget för grupparbetet, samt val av bedömningskriterier och implementering av dessa granskades. Utefter detta underlag arbetades bedömningskriterier samt ett system för inbördes poängsättning och efterföljande individuell betygsättning av grupparbetet fram. Metoden testades på en grupp med studenter och resultaten visar att inbördes bedömning kan fungera bra och att majoriteten av studenterna upplevde bedömningsprocessen övervägande positivt.

Nyckelord

Inbördes bedömning, peer evaluation, betygsättning, examination, grupparbeten

Bakgrund

Inom mitt ämnesområde, marin ekologi, är det mycket vanligt med grupparbeten i undervisningen. Själv undervisar jag på en kurs på masternivå 15 hp (BIO451, Experimentell marin ekologi) där ett experimentellt grupparbete avslutar kursen och upptar ca 60 % av totala kurstiden. Studenterna jobbar i grupper om 3-5 personer och kursen ges på universitetets marina forskningsstation på Tjärnö utanför Strömstad där studenterna bor under kurstiden. Grupparbetet avslutas med inlämnande av en gemensam skriftlig rapport och en muntlig presentation. För närvarande så betygsätts inte grupparbetet annat än med godkänt och närvaro under hela grupparbetestiden är obligatorisk. Studenternas slutbetyg på kursen (U, G eller VG) avgörs helt av resultatet på den skriftliga tentamen som ges innan grupparbetesdelen börjar. En del studenter har uttryckt sitt missnöje över och tycker att det är orättvist att grupparbetet inte betygsätts när det nu utgör en så stor del av kursen. Eftersom grupparbetet vid sidan av en muntlig och skriftlig presentation också innehåller praktiska moment som planering och utförande av experiment för att angripa en ekologisk frågeställning och påföljande analys av resultat, är dilemmat hur man som lärare ska kunna bedöma enskilda studenters insatser inom alla delar av grupparbetet. Med bakgrund av detta vill jag undersöka hur sådana här grupparbeten kan bedömas för att på bästa sätt stimulera studenter att aktivt delta i alla delar av arbetet, att studenterna lär sig så mycket som möjligt och att enskilda studenter får en rättvis bedömning och betygsättning. Jag har valt att titta närmare på och använda mig av metoden där studenterna utvärderar (betygsätter) varandra genom inbördes bedömning. Mer specifikt så vill jag i detta arbete:

- Göra en djupdykning i litteraturen som handlar om grupparbeten, betygsättning och särskilt metoder där studenter inbördes bedömer varandra
- Utarbeta ett formulär med kriterier för studenternas bedömning av varandra och anvisningar hur de ska poängsätta varandra
- Ta fram en modell för hur studenternas bedömning av varandra ska implementeras i betygsättningen
- Testa inbördes bedömning på en grupp studenter som gör ett experimentellt grupparbete och försöka analysera hur utvärderingen gick samt att ta fram ett frågeformulär om hur de upplevde bedömningsprocessen och sammanfatta resultatet.

Grupparbeten och betygsättning

Grupparbeten är vanligt förekommande på universitet och högskolor. Genom grupparbeten så utvecklar studenterna personliga färdigheter som kan användas även utanför deras egna specifika ämnesområden och som kan appliceras i många olika sammanhang och situationer (Shah, 2013). Exempel på sådana personliga överförbara färdigheter är kommunikation, lagarbete, analytisk kapacitet och förmåga att lösa problem. Dessa färdigheter anses mycket värdefulla och är högt efterfrågade av arbetsgivare (Shah, 2013). Grupparbeten är också viktiga för studenternas utveckling av ämnesspecifika förmågor, djupare tänkande och för ett aktivt lärande (Johnston & Miles, 2004; Plastow, Spiliotopoulou & Prior 2010; Shah, 2013). Flera studier har visat att studenter är mycket nöjda med grupparbeten (Gatfield, 1999; Knight, 2004).

Det finns också ett flertal nackdelar med grupparbeten. Från studentens perspektiv så inkluderar dessa: brist på engagemang, dålig organisation, gruppdynamik innefattande konflikter och dålig gruppssammanhållning, varierande nivåer av färdigheter och kunskaper inom gruppen, samt att vissa studenter glider med eller åker snålskjuts på andra (Plastow et al., 2010 och referenser däri). Problemet med gruppmedarbetare som själva bidrar väldigt lite och därmed drar nytta av andras arbete är ett vanligt klagomål bland studenter vad gäller otillfredsställande erfarenheter av grupparbete (Brooks & Ammons, 2003). Ur ett lärarperspektiv så är frågan hur grupparbetena ska bedömas den mest diskuterade och problematiska (Plastow et al., 2010).

Metoder att betygsätta grupparbeten

Bedömning av grupparbeten sker vanligen genom att gruppen lämnar in en gemensam projektrapport och eftersom läraren inte har möjlighet att följa hela arbetet i gruppen så är enskilda studenters inläring och insatser inte kända (Gammie & Matson, 2007 och referenser däri).

Race (2000) (i Gammie & Matson, 2007) listar sju möjliga tillvägagångssätt att bedöma grupparbeten:

- I. Ge alla i gruppen samma betyg
- II. Dela upp arbetet mellan gruppmedlemmarna och betygsätt varje del separat
- III. Betygsätt grupparbetet som helhet och förhandla om skillnader mellan gruppmedlemmar
- IV. Betygsätt grupparbetet som helhet och be gruppmedlemmarna att inbördes bedöma varandras insatser som ger ytterligare betyg
- V. Betygsätt grupparbetet som helhet och lägg till individuellt bedömda uppgifter
- VI. Betygsätt grupparbetet som helhet och lägg till individuell muntlig examination
- VII. Betygsätt grupparbetet som helhet och bedöm individuella skillnader genom specifika frågor i examinationen/på tentan som handlar om grupparbetet

Andra varianter som jag hört talas om eller stött på är att gruppmedlemmarna skriver arbetet i ett gemensamt dokument i/på en webbaserad plattform (ex GUL) så att läraren kan se vem som bidrar med vad eller för experimentella grupparbeten att studenterna skriver material och metod- och resultat-delen tillsammans men inledning och diskussion individuellt. Chapman & Van Auken (2001) konstaterar att om metoder för att bedöma individuell prestation tillämpas är studenterna mer positiva till grupparbeten. En vanlig form för att utvärdera grupparbeten verkar vara att använda sig av inbördes bedömning (peer evaluation) där varje gruppmedlem bedömer det individuella bidraget för var och en av gruppmedlemmarna (Goldfinch & Raeside, 1990; Goldfinch, 1994; Mathews, 1994; Gatfield, 1999; Li, 2001; Brooks & Ammons, 2003; Johnston & Miles, 2004; Kennedy, 2006; Gammie & Matson, 2007). I sammanfattningen av Race (2000) ovan så ingår inbördes bedömning i alternativ III och IV.

Inbördes bedömning

Inbördes bedömning innebär att studenterna får betygsätta varandras bidrag till grupparbetet. I vissa fall förväntas studenterna också utvärdera sin egen insats (t.ex. Goldfinch, 1994; Johnston & Miles, 2004). Studenterna producerar vanligen en gemensam skriftlig rapport (t.ex. Mathews, 1994; Gammie & Matson, 2007) men inbördes bedömning kan också tillämpas tillsammans med individuellt skriven rapport (Johnston & Miles, 2004). Betygsättningen på grupparbetet består av 2 delmoment: En del av betyget sätts av läraren som bedömer den skriftliga rapporten (och/eller eventuell annan slutprodukt); den andra delen bestäms av gruppmedlemmarna. I en vanligt använd modell så resulterar gruppmedlemmarnas bedömning i en s.k. bidragsfaktor som används för att justera individens betygspoäng relativt till det betygspoäng som den skriftliga rapporten fick (t.ex. Li, 2001). I en annan modell så används gruppmedlemmarnas bedömning som helt separata betygspoäng som adderas till rapportens betygspoäng. I många fall ombeds studenterna att poängsätta olika delar och aspekter av gruppmedlemmarnas insatser (Abelson & Babcock, 1986; Goldfinch & Raeside, 1990; Goldfinch, 1994; Mathews, 1994; Gatfield 1999; Li, 2001; Gueldenzoph & May, 2002; Gammie & Matson, 2007) men ibland är det bara en allomfattande bedömning som ska ges (t.ex. Kennedy 2006).

Metoder för inbördes poängsättning i grupparbetet

Det har genom åren provats flera olika poängsättnings-metoder för inbördes bedömning. Vanligen får alltså studenterna utvärdera varandra i flera olika kategorier. I vissa fall har man i alla fall delvis använt sig av "ett rankingsystem" där man fått fylla i namnet på den person som bidragit mest till en specifik del. Denna ranking har sedan omvandlats till poäng av läraren. Vad

gäller poängsystem så kan dessa delas upp i 2 huvudgrupper. I det ena huvudsakliga systemet så sätter studenterna poäng efter en relativ skala (Abelson & Babcock, 1986; Goldfinch & Raeside, 1990; Goldfinch, 1994; Gatfield, 1999; Li, 2001; Gueldenzoph & May, 2002) exempelvis 1-5 där 1 motsvarar minimalt bidrag, 3 motsvarar medelbidrag och 5 motsvarar utomordentligt bidrag. Det finns varianter där man inkluderar 0 (bidrag inte alls) och t.o.m. -1 (motverkade grupparbetet) (Goldfinch, 1994; Gatfield, 1999). I tidiga försök användes poängsumman direkt som en bidragspoäng (Goldfinch & Raeside, 1990) men ganska snart identifierade man då problemet att i högpresterande grupper blev medelpoängen vid den inbördes bedömningen högre vilket ledde till att dessa gruppmedlemmar "belönades" 2 ggr och tvärtom för medlemmar i de grupper som presterade lägre som "bestraffades" 2 ggr. Därför provade man istället att dividera individens poäng med medelpoängen för gruppen för att få fram en bidragsfaktor. För att omvandla poängen som en individ fått till en bidragsfaktor så divideras individens poäng med medelpoängen för gruppen.

I den andra huvudsakliga modellen får studenterna en viss konstant poängsumma att fördela ut bland gruppmedlemmarna (Mathews, 1994; Kennedy, 2006). Exempelvis kan studenterna få 100 poäng totalt att fördela ut (Mathews, 1994). En annan variant är att studenterna får en total poängsumma att dela ut som beror på gruppstorleken eller antalet gruppmedlemmar som bedöms. Mathews (1994) testade en modell där varje student fick $10*n$ poäng att dela ut där n är antalet gruppmedlemmar som utvärderas. Enligt Kennedy (2006) är en vanlig modell på australiska universitet ett poängsystem med $100*n$ poäng. En individuell bidragsfaktor fås här fram mer direkt genom att man räknar ut medelpoängen för varje student och i det senare fallet fås bidragsfaktorn i % direkt.

Det finns en viktig skillnad mellan de 2 huvudsakliga poängsystemen och det är att i det första fallet med relativa poäng, även om bidragsfaktor räknas ut, fås skillnader i den totalpoäng som varje enskild individ delar ut. Det leder till att de studenter som är generösare än sina kamrater och i medeltal ger högre poäng själva blir bestraffade (Goldfinch, 1994). I Goldfinch (1994) prövades att införa självbedömning för att komma runt problemet. Tanken var att studenterna skulle vara lika generösa mot sig själva som mot andra. Men Johnston & Miles (2004) konstaterade i sin undersökning att studenter generellt värderar sin egen insats högre än vad medstudenterna gör och speciellt gjorde de högpresterande studenterna det. Detta gällde dock inte de studenter vars prestationsförmåga låg i den lägre kvartilen som istället undervärderade sin egen insats. Ett annat sätt komma runt problemet testades av Li (2001) som normaliserade den utdelade poängen till medelpoängen som den enskilda studenten gett. Denna åtgärd leder i

praktiken till samma resultat som när en konstant poängsumma används! Båda poängsystemen har dock en kvarvarande svaghet som identifierades av Bushell (2006) nämligen att rangordningen mellan gruppmedlemmarna ändå kan bli felaktig (när självbedömning inte används) vilket inträffar då den mest kvalificerade studenten i gruppen rankar den näst mest kvalificerade mycket högre än de övriga gruppmedlemmarna. Man bör alltså alltid kolla så att rangordningen och totala poängsättningen överensstämmer och om inte så presenterar Bushell (2006) en metod att lösa problemet.

Hur resultatet av inbördes bedömning implementeras i totala betyget för grupparbetet

Det har också provats olika metoder för hur inbördes bedömning implementeras i betyget för grupparbetet. Jag väljer här att fördjupa mig i den modell där gruppmedlemmarnas bedömning används för att justera individens betygspoäng uppåt eller nedåt relativt till betyget som den gemensamma grupprapporten fått av läraren. Det mest direkta sättet är att använda den framräknade individuella bidragsfaktorn och multiplicera med grupp-betyget. Om en students bidrag genom inbördes bedömning värderas till i genomsnitt 90 % så multipliceras betygspoängen för grupparbetet med 0.9 för att få fram individens slutbetyg för grupparbetet. Det har visat sig att en direkt användning av bidragsfaktorn på detta sätt kan leda till extremt stora skillnader i betygspoäng inom en grupp (Goldfinch & Raeside, 1990). Därför har varierande viktning av studenternas bedömning av varandra och lärarens betyg för grupparbetet testats. Exempelvis kan hälften (alternativt 1/3) av grupp-betygspoängen behållas intakt av varje student medan andra hälften (alternativt 2/3) av betygspoängen multipliceras med den individuella bidragsfaktorn (Goldfinch & Raeside, 1990). En annan variant som leder till att studenter som bidrog väldigt lite får ett rättvisare betyg är att använda sig av formeln:

Studentens betygspoäng = (Individuell bidragsfaktor)^{0.5} * Grupp-betygspoängen

(Goldfinch, 1994). Här kan exponenten varieras (0-1) för mindre eller större del av betyget som bestäms av individuella bidragsfaktorn (egen anmärkning).

Utformning och tillämpning av bedömningskriterier

Jag har inte hittat så mycket i litteraturen där betydelsen av utformningen av bedömningsfrågorna undersökts. Kennedy (2006) visar dock att inbördes bedömning där man bara ger en sammanlagd poäng och inte betygsätter flera bestämda kriterier leder till stora individuella skillnader i hur man poängsätter en specifik gruppmedlems bidrag. Hans tolkning är att studenterna då inte riktigt vet vad de ska utvärdera hos varandra. Även Brooks & Ammons (2003) konstaterar att specifika

bedömningskriterier minskar problemen vid inbördes bedömning. Matthews (1994) delade in kriterierna för bedömningen i 3 områden:

- *Projektets output* - bidrag till uppgiften i form av tid/ansträngning/kvalitet
- *Uppgiftshantering* - handlingsätt som leder gruppen framåt mot uppgiftens mål
- *Processhantering* - handlingsätt som bestämmer kvalitén på relationerna gruppmedlemmarna emellan

De två sistnämnda områdena anses av beteendevetare representera centrala dimensioner vad gäller gruppprocesser och är därmed viktiga vid att ha med vid bedömning (Matthews, 1994). Denne författare betonar också vikten av att studenterna behöver få tillgång till bedömningskriterierna vid grupparbetets start för att kunna följa och bedöma medarbetarnas prestationer.

Material och metod

Med stöd av litteraturen arbetade jag fram ett bedömningsformulär med ett antal kriterier för inbördes bedömning, och med instruktioner för poängsättning. Målet med kriteriernas utformning var att individernas bidrag till både grupprelaterade processer och produktens färdigställande skulle utvärderas. Detta innefattar bl. a. studenternas förmåga att samarbeta och driva grupparbetet framåt, kontroll av deras generella engagemang och närvaro, samt utvärdering av till hur stor del de bidrog till att de olika delarna av arbetet blev färdigställda och lyckade. Vid utformning av kriterierna tänkte jag också på att de även skulle fungera formativt. d.v.s. att de skulle stimulera studenterna till att aktivt delta i alla delar av arbetet. Jag har hämtat inspiration till bedömningskriterierna i följande referenser: Abelson & Babcock (1986), Goldfinch & Raeside (1990), Goldfinch (1994), Mathews (1994), Gatfield (1999), Li (2001), Gueldenzoph & May (2002), Gammie & Matson (2007). Vad gäller poängsättningen var min tanke att den ska passa för gruppstorlekar på 3-5 studenter. Vidare ville jag att den skulle vara enkel att förstå, ge liknande resultat oberoende av gruppstorlek och individuell "generositet" gentemot gruppkamraterna. Genom att granska befintliga modeller i litteraturen arbetade jag också fram ett förslag till modell för hur resultatet av den inbördes bedömningen ska implementeras i studentens slutbetyg för grupparbetet.

Slutligen testade jag också bedömningsprocessen på en grupp med studenter. Mina teststudenter gick andra året på marina programmet och deltog på kursen i marin hårdbottensekologi, MAR108, under hösten 2015. De 21 studenterna gjorde under sin fältvecka på den marina fältstationen på Tjärnö ett 2.5-dagars experimentellt grupparbete och fick på frivillig basis fylla i blanketten för

inbördes bedömning där självbedömning ingick (bilaga 2). Studenterna uppmanades att läsa igenom bedömningskriterierna innan grupparbetet startade. Bedömningen skedde bara på prov och varje student hittade på ett fingerat namn som alla i gruppen använde vid utvärderingen av varandra. Studenterna jobbade i 7 grupper om vardera 3 studenter. Arton av 21 studenter fyllde helt eller delvis i blanketten, sista kriteriet på listan (skriva rapporten) ströks eftersom studenterna inte skulle skriva en gemensam rapport. Totalt utvärderades alltså 13 olika kriterier. Vid uppföljningen av fältveckan bad jag också studenterna fylla i en blankett om hur de upplevde bedömningsprocessen (bilaga 3) och även här lämnades 18 blanketter in.

Resultat

Framtagning av bedömningskriterier samt system för inbördes poängsättning och betygsättning

Jag tog fram två varianter av bedömningsformuläret, ett där självvärdering ingår och ett där bara de övriga gruppmedlemmarna utvärderas (bilaga 1 och 2). Bedömningen utformades så att studenterna får fördela ut 100 poäng (motsvarande %) för varje bedömningskriterium/delmoment som ska avspegla hur stor del de anser att respektive student bidragit med. Studentens genomsnittliga poäng (medel beräknat över alla kriterier och från alla bedömare) delas sedan med medel för det utvärderade antalet studenter (33.3 om 3 studenter, 25 om 4 studenter o.s.v.). Ett resultat nära 1 betyder att den enskilda studentens bidrag är medel, under 1 innebär att studenten har underpresterat i gruppen och ett resultat över 1 innebär att studenten har bidragit med mera än medeltalet. Observera att lägsta resultat är 0 men högsta möjliga resultat beror på gruppstorleken. Det inlämnade grupparbetet poängsätts för hela gruppen gemensamt och sedan kan denna poäng multipliceras med den individuella bidragsfaktorn för att ge betygspoäng för varje enskild student för grupparbetet. Hurvida den individuella bidragsfaktorn kan användas direkt eller bara delvis på gruppabetespoängen får testas fram. Om direkt multiplicering med bidragsfaktorn på hela gruppabetespoängen genomgående ger mycket stora betygsskillnader inom grupper kommer jag prova att använda bara en del av bidragsfaktorn eller bara multiplicera bidragsfaktorn med en del av gruppabetet. Det inlämnade grupparbetet behöver betygsättas i en skala på max 100 p som relaterar till de betygsgränser som gäller för kursen (60% för G och 85% för VG). Exempelvis kan ett grupparbete få 95 p om det är riktigt bra men enskilda individer i gruppen kan halka under 85 % om man får en individuell bidragsfaktor under 1. Notera att en student också kan få högre sammanlagt betyg än 100 på grupparbetet.

Test av bedömningsprocessen på studenterna

Inbördes bedömning

De 18 inlämnade bedömningsblanketterna var fördelade så att i fyra grupper så fyllde alla tre studenterna i blanketten och i de övriga tre grupperna så fyllde två medlemmar i blanketten. För många av kriterierna i många grupper så gjordes ingen åtskillnad mellan gruppmedlemmarna då alla gavs samma poäng. Detta hände i medeltal i 71 % av kriteriefallen men skilde sig stort mellan grupperna (min 45 % max 97 %), se vidare diskussionen. Då medelpoängen över alla kriterier som en enskild student givit till en annan omräknas till bidragsfaktor så hamnar den lägsta faktorn på 0.83 och den högsta på 1.11. I grupper där 2 personer utvärderat en tredje och ett medel räknas fram från de båda bedömningarna ligger lägsta bidragsfaktorn på 0.87 och högsta på 1.07. Motsvarande bidragsfaktorer som man givit sig själv ligger mellan 0.92 och 1.10. Åtta studenter värderade i medeltal sig själva lägre än vad kompisarna gjorde och 10 värderade sig själva högre. Om man ser till avvikelser i grupperna där alla 3 medlemmarna deltog i bedömningen så värderar den student som sticker ut mest sig själv i medeltal 10.6 % högre än de andra gruppmedlemmarna gör och i andra änden är det två studenter som värderar sig själva 3.5 % lägre än deras gruppkompisar gör. Men studenterna är inte generellt "sämre" på att bedöma sin egen insats jämfört med hur kompisarna utvärderar dem. Standardavvikelsen i betygsfaktor räknades ut då betygsfaktorerna för varje enskild student parades ihop på följande sätt: a) själv-faktor från kompis 1; b) själv-faktor från kompis 2; c) faktor från kompis 1-faktor från kompis 2. Medel för standardavvikelsen var i a) 0.027, i b) 0.030, och i c) 0.037 vilket betyder att bidragsfaktorerna som kompisarna gav en student i medeltal skiljer sig mer åt än vad bidragsfaktorn man gav sig själv jämfört med faktorn man fick av kompisarna gjorde.

Upplevelse av bedömningsprocessen

Flertalet studenter hade aldrig gjort inbördes bedömning och av de 5 som gjort det hade en gjort det i gymnasiet och en bara muntligt. Vad gäller frågorna hur de upplevde att utvärdera de andra i gruppen respektive sig själva så svarade knappt hälften av studenterna att de tyckte det var svårt att bedöma de andra i gruppen och svårt eller ännu svårare att utvärdera sig själva. Övriga studenter tyckte att det var OK, bra, positivt, känsligt eller lärorikt. Av kommentarerna att döma tyckte flera av de som upplevde det besvärligt att det var svårt att särskilja studenternas insatser p.g.a. att alla i gruppen jobbat lika mycket, grupparbetet var så kort, samt att det var naturligt under omständigheterna att jobba lika mycket. Svårigheten i att bedöma sin egen insats låg främst i osäkerhet om att vara självkritisk/återhållsam eller självgod/överdriven och en student skrev att

hen aldrig gav sig själv högre poäng än de andra. De som tyckte det var positivt att utvärdera sig själva uttrycker att de tror att det är utvecklande och att man tydligare ser sin egen insats i grupparbetets olika delar. Angående poängsystemet tyckte en student att det var svårt att göra bedömningen i procentsatser och en att det var konstigt att utvärdera i siffror istället för med ord. En tredje tycker att en skala på t.ex. 0-5 hade varit bättre så man inte behöver välja hur man fördelar. Ytterligare två studenter efterlyser möjlighet att skriva en motivering till sin bedömning för att det ska fungera mer konstruktivt.

Mer än hälften av studenterna anser att deras arbete i gruppen påverkades positivt av att bedömningskriterierna delades ut vid starten av grupparbetet. De bedömer att det ledde/kan leda till ökat bidrag och engagemang från alla gruppmedlemmar och att det motiverar gruppmedlemmarna att vara delaktig i alla moment. En student påpekade att det var bra med de olika kriterierna för då kunde man se vad man kunde göra under grupparbetet och en annan att hen reflekterade över gruppens arbetsgång mer. Kritiska röster befärar att denna typ av bedömningsprocess kan leda till att folk tar mer för sig (jag tolkade detta som negativt) och att det kan ta fokus från själva arbetet. Ett par röster tycker inte att det påverkade alls eftersom bedömningen skedde i efterhand eller att de inte hade läst kriterierna innan.

På frågan om studenterna tycker att individuell betygsättning av grupparbeten är viktigt/mer rättvist så var det inte någon som svarade rent nej medan 6 studenter var vankelmodiga och 12 svarade ja. Bland de tveksamma hörs röster om att ett grupparbete bör få vara just ett grupparbete, men i ett fall villkorat med att det finns något annat sätt att visa sin individuella kunskap på. En student tycker det är mer viktigt att se till att det finns dialog inom gruppen så att det kan vara en bra arbetsfördelning. Det finns en oro över att bli orättvist bedömd av de andra, men flertalet av de tveksamma uttrycker att det är viktigt att någon som bidrar väldigt lite inte ska kunna bli godkänd. De positiva studenterna motiverar sitt ja med att det är viktigt att inte någon åker snålskjuts och att de som är skickligare och som gör större insatser också belönas. Vad gäller användning av inbördes bedömning som metod för individuell betygsättning så svarar bara en student att hen inte tycker det är bra. Motivationen är risken för orättvis bedömning om några i gruppen är bra kompisar eller mobbar ut en speciell medlem. Fyra studenter är tveksamma varav problem med att man kan "fuska" och skriva att man gjorde mer än vad man egentligen gjorde och att det kan vara svårt att bedöma dem man känner nämns. En student anser också att metoden inte leder till att man utvecklas eftersom man bedöms "bakom ryggen" och inte får positiv eller negativ kritik uttryckt i ord. Tolv av studenterna är generellt positiva men återkommande kommentarer tyder på oro att inte bedömningen kan ske helt ärligt, objektivt och utan förutfattade

meningar, speciellt när gruppmedlemmarna känner varandra sen tidigare. Oron gäller både att inte kunna ge och risken att inte få en rättvis bedömning. Flera studenter uttrycker att inbördes bedömning kan ge läraren en bra bedömningsgrund för betygsättning och att bedömningen gör att man individuellt skärper sig.

Som alternativ/komplement till skriftlig inbördes bedömning med poäng kom studenterna med följande förslag: a) att alla skriver egen rapport, b) att projektet övervakas av lärare och studenterna inbördes namnger vem som gjort vad, c) heta stolen med konstruktiv kritik, samt d) mer muntlig utvärdering. Vad gäller kriterierna som bedömdes hade studenterna inget förslag på något kriterium de tyckte fattades och flera uttryckte att de täckte in det mesta.

Diskussion

Inbördes bedömning framträder i denna studie som ett bra alternativ för individuell betygsättning av grupparbete. Jämfört med exempelvis inlämning av egen rapport efter grupparbetet så föredrar jag modellen där studenterna vid grupparbetets slut lämnar in en gemensam rapport eftersom det bäst avspeglar verkligheten då man jobbar i en projektgrupp. Det är ju en gemensam slutprodukt vilken alla har bidragit till som gäller och man behöver komma överens om innehåll och utformning. Rapporten tycker jag därmed ska betygsättas för hela gruppen gemensamt och sedan är det rapportens poäng som ligger till grund för studentens individuella betyg genom att individen får mer eller mindre än gruppens betygs-poäng beroende på prestation. Jag anser också att detta är ett rättvisare system för det är ju så mycket mer än skriva och sammanställa som ingår i ett framgångsrikt grupparbete. T. ex kan vissa studenter vara duktiga på att organisera grupparbetet, lösa problem, tolka data, praktiskt utföra experiment mm som alla andra i gruppen drar nytta av. En viktig fråga är om studenterna även ska bedöma sin egen prestation i gruppen eller bara de andras. Båda modellerna finns (t.ex. Goldfinch, 1994; Kennedy, 2006) och efter utvärderingen i denna studie har min ursprungliga utgångspunkt stärkts vilken var att självbedömning inte bör ingå. Mina teststudenter övervärderade visserligen inte generellt sig själva vilket Johnston & Miles (2004) fann men den stora osäkerheten studenterna kände inför att utvärdera sig själva tycker jag är talande.

Att bedömningskriterierna utdelas vid grupparbetets start är viktigt för att studenterna ska kunna bedöma varandra kontinuerligt under hela processen (Matthews, 1994). Detta tycks också ha fungerat formativt vilket var avsikten; mina teststudenter upplevde att bedömningsprocessen stimulerade gruppmedlemmarna till ökat engagemang och motiverade till delaktighet i alla moment. Chapman & Van Auken (2001) fann att tidig implementering av inbördes bedömning

tillsammans med ett flertal utvärderingar under projektets gång minskade problemen med passiva studenter och generellt ökade nivån av tillfredsställelse med grupparbetet bland studenterna. I det 6 veckor långa grupparbetet på BIO451 är det troligen bra att ha en delbedömning ett par veckor in i arbetet. Majoriteten av studenterna var också positiva till individuell betygsättning av grupparbete. Även tidigare undersökningar har visat att om det finns ett väl fungerande system för bedömning av individuella prestationer så är studenter mer positiva till grupparbeten och föredrar summativ framför formativ utvärdering (Chapman & Van Auken, 2001; Shah, 2013).

Vad gäller bedömningskriterierna så var jag rädd att de skulle upplevas som för många men ingen av studenterna kommenterade det. Alla studenter hade inte fyllt i alla kriterier men det kan bero på att grupparbetet var så kort och att alla kriterier inte riktigt var applicerbara (exempelvis "leverera bidrag på utsatt tid"). Jag kommer dock att se över bedömningskriterierna efter synpunkter från andra kursdeltagare på HPE103 som påpekade att vissa bedömningskriterier kräver att studenterna som gör bedömningen själva har en viss grad av kunskap och förståelse; detta gäller exempelvis kriterierna "Förståelse för uppgiften" och "Sätta in resultaten i ett sammanhang". Poängsättningssystemet tycks ha fungerat ok, men några studenter föreslår andra varianter eller komplement. Den föreslagna varianten med 0-5 p är ju en relativ skala med alla problem och omräkningsfaktorer som krävs (se under "Metoder för inbördes poängsättning i grupparbetet"). Jag vill hålla fast vid varianten med en konstant poängsumma men kan tänka mig att som alternativ använda $100 * n$ poäng att fördela där n är antalet gruppmedlemmar som bedöms (Kennedy, 2006). Jag tycker att denna variant känns mer gruppstorleksberoende, d.v.s. att det kanske är svårare att få exempelvis hög bidragsfaktor i en stor grupp än i en mindre. Metoden jag här använt med 100 p att fördela känns dock inte användbar om fler än 5 gruppmedlemmar ska utvärderas. Jag tycker däremot att komplettera sifferbedömningen med skriftliga kommentarer är ett väldigt bra förslag som jag vill lägga till på bedömningsblanketten. Dessa kommentarer kommer inte användas för betygsättningen men kan ges till respektive student som då kan få konstruktiv kritik som kan leda till utveckling.

Den stora andelen kriteriefall där gruppmedlemmarna inte gjort någon åtskillnad mellan varandra utan givit alla samma poäng kan enligt kommentarer från studenterna delvis förklaras med att grupparbetet var kort och därmed svårbedömt men berodde också i många fall på att studenterna verkligen jobbade lika mycket (både kvalitativt och kvantitativt) med alla delar. Det tyder ju på att gruppen och deras arbete fungerat väldigt väl och då bör resultatet inte heller bli några stora skillnader i bidragsfaktor. Gammie & Matson (2007) konstaterar dock att det finns en motvillighet att göra åtskillnad mellan sina gruppkamrater. Utvärderingen i denna studie

resulterade i individuella bidragsfaktorer mellan 0.87 och 1.07 i fall där 2 personer bedömt en tredje. Om denna bidragsfaktor används direkt för att räkna ut betygspoängen innebär det en sänkning av individuella betyget gentemot gruppbetyget med 13 % respektive en höjning med 7 %. Detta spann verkar inte orimligt stort men ett mer långvarigt grupparbete kanske leder till betydligt större åtskillnad och en alternativ viktning av bidragsfaktorn då bör användas. En bidragsfaktor på 0.87 innebär trots allt att studenten hamnar under gränsen för godkänt (60%) om grupparbetet får 68 poäng eller lägre.

Slutsats

För att summera ser jag inbördes bedömning som ett fullt användbart redskap att betygsgöra individer i grupparbeten. Både mina teststudenters positiva attityd och resultaten från bedömningsblanketterna tyder på det. Jag ser gärna att inbördes bedömning testas på masterkursen i Experimentell marin ekologi under våren 2016. För att ta det säkra före det osäkra kanske systemet skulle provas en omgång även på denna kurs innan det eventuellt används på allvar. Genom att använda individuell bedömning kan rättvisare betygssättning uppnås. I tillägg verkar bedömningsprocessen kunna fungera formativt genom att stimulera till ökad läraktivitet vilket sannolikt resulterar i ökad inläring och att lärandemålen uppfylls i högre grad.

Tack

Varmt tack till studenterna på MAR108 HT 15 för deras deltagande som teststudenter i denna studie.

References

- Abelson, M.A., & Babcock, J.A. (1986). Peer evaluation within group projects: a suggested mechanism and process. *Journal of Management Education*,
- Brooks, C.M., & Ammons, J.L. (2003). Free riding in group projects and the effects of timing, frequency, and specificity of criteria in peer assessments. *Journal of Education for Business*, 78(5), 268–272.
- Bushell G. (2006). Moderation of peer assessment in group projects. *Assessment and Evaluation in Higher Education*, 31(1), 91–108.
- Chapman, K.J., & Van Auken, S. (2001). Creating Positive Group Project Experiences: An Examination of the Role of the Instructor on Students' Perceptions of Group Projects. *Journal of Marketing Education*, 23(2), 117-127.

- Gammie, E., & Matson, M. (2007). Group assessment at final degree level: An evaluation. *Accounting Education*, 16(2), 185–206.
- Gatfield, T. (1999). Examining student satisfaction with group projects and peer assessment. *Assessment and Evaluation in Higher Education*, 24(4), 365–377.
- Goldfinch, J. & Raeside, S. (1990). Development of a peer assessment technique for obtaining individual marks on a group project. *Assessment and Evaluation in Higher Education*, 15(3), 210–231.
- Goldfinch, J. (1994). Further developments in peer assessment of group projects. *Assessment and Evaluation in Higher Education*, 19(1), 29-35.
- Gueldenzoph, L.E. & May, G.L. (2002). Collaborative peer evaluation: best practices for group member assessment. *Business Communication Quarterly*, 65(1), 9-20.
- Johnston, L., & Miles, L. (2004). Assessing contributions to group assignments. *Assessment and Evaluation in Higher Education*, 29(6), 751–768.
- Kennedy, G.J. (2006). Peer-assessment in group projects: is it worth it? *Conferences in Research and Practice in Information Technology*, 42.
- Knight, J. (2004). Comparison of student perception and performance in individual and group assessments in practical classes. *Journal of Geography in Higher Education*, 28(1), 63–81.
- Li, L.K.Y. (2001). Some refinements on peer assessments of group projects. *Assessment and Evaluation in Higher Education*, 26(1), 5–18.
- Mathews, B.P. (1994). Assessing individual contributions: experience of peer evaluation in major group projects. *British Journal of Educational Technology*, 25(1), 19-28.
- Plastow, N., Spiliotopoulou, G., & Prior, S. (2010) Group assessment at first year and final degree level: a comparative evaluation. *Innovations in Education and Teaching International*, 47(4), 393–403.
- Race, P. (2000) *Tips on Group Learning* (London: Kogan Page).
- Shah, S.Z.A. (2013). The use of group activities in developing personal transferable skills. *Innovations in Education and Teaching International*, 50(3), 297-307.

Grupp-projekt

Inbördes bedömning (exklusive självbedömning)

Namn _____ Datum _____

Kurs _____ Antal i gruppen _____

Projektets titel _____

Instruktioner: Var god utvärdera alla medlemmar i din grupp exklusive dig själv. Du ska bedöma gruppmedlemmar enligt vad du tycker är deras respektive bidrag (både kvalitativt och kvantitativt) till ert arbete i kategorierna nedan. Du ska tänka på till hur stor del de olika gruppmedlemmarna bidrog på så sätt att summan i varje kategori alltid blir 100 %. Du ska alltså helt bortse från din egen insats. Om du t.ex. utvärderar 3 andra gruppmedlemmar och tycker att de presterat helt lika i en kategori skriver du 33.3 i alla medlemmars rutor.

	För och efternamn på gruppdeltagare				
Bedömningskategori					
Närvaro under grupparbetet					
Bidrag till gruppkänsla, förmåga /vilja att inkludera alla i arbetet					
Leverera bidrag till arbetet på utsatt tid (möta deadlines)					
Organisera och föra grupparbetet framåt					
Gjorde sin beskärda del av totala uppgiften					
Konstruktiva idéer som utvecklade arbetet					
Lösa uppkomna problem					
Förståelse för uppgiften, insamling av bakgrundsfakta					

Design av experiment för att angripa frågeställningen					
Utföra den experimentella uppgiften					
Analys och tolkning av data					
Presentera resultat (tabeller, diagram)					
Sätta in resultaten i ett sammanhang					
Skriva rapporten					

Grupp-projekt

Inbördes bedömning (inklusive självbedömning)

Namn _____ Datum _____

Kurs _____ Antal i gruppen _____

Projektets titel _____

Instruktioner: Var god utvärdera alla medlemmar i din grupp inklusive dig själv. Du ska bedöma varje gruppmedlem enligt vad du tycker är ditt/deras respektive bidrag (både kvalitativt och kvantitativt) till ert arbete i kategorierna nedan. Du ska tänka på till hur stor del de olika gruppmedlemmarna bidrog på så sätt att summan i varje kategori alltid blir 100 %. Du ska även bedöma din egen insats. Om ni t.ex. är 3 medlemmar i gruppen och du tycker att ni presterat helt lika i en kategori skriver du 33.3 i alla medlemmars rutor.

	För och efternamn på gruppdeltagare				
Bedömningskategori	Jag själv				
Närvaro under grupparbetet					
Bidrag till gruppkänsla, förmåga /vilja att inkludera alla i arbetet					
Leverera bidrag till arbetet på utsatt tid (möta deadlines)					
Organisera och föra grupparbetet framåt					
Gjorde sin beskärda del av totala uppgiften					
Konstruktiva idéer som utvecklade arbetet					
Lösa uppkomna problem					
Förståelse för uppgiften, insamling av bakgrundsfakta					

Design av experiment för att angripa frågeställningen					
Utföra den experimentella uppgiften					
Analys och tolkning av data					
Presentera resultat (tabeller, diagram)					
Sätta in resultaten i ett sammanhang					
Skriva rapporten					

Utvärdering av inbördes bedömning i grupp-projekt

Dessa frågor delades ut till mina test-studenter som på prov bedömde varandra under ett experimentellt grupp-projekt. Det kan också tänka sig att några av dessa frågor kommer att ingå i kursutvärderingen på kurser där inbördes bedömning kommer att tillämpas.

Har du utfört inbördes bedömning någon gång förut?

Hur upplevde du att utvärdera de andra i gruppen?

Hur upplevde du att utvärdera dig själv?

Hur tror du att det påverkade grupparbetets gång att bedömningsfrågorna delades ut vid starten?

Tycker du att individuell betygsättning av grupparbeten är viktigt/mer rättvist?

Tycker du att inbördes bedömning är ett bra system för detta?

Om inte, kan du tänka på något annat sätt?

Någon/några viktiga kriterier som du tycker fattas på blanketten för inbördes bedömning?