NORGES HANDELSHØYSKOLE

Bergen, Våren 2017

Pay Enough or Don't Rate at All

En eksplorerende casestudie av prestasjonsevalueringer basert på skjønn og tvungen karakterfordeling

Birgitte Kielland og Idunn Marie Svendsen

Veileder: Iver Bragelien

Masterutredning i hovedprofilene Strategi og Ledelse og Økonomisk Styring

NORGES HANDELSHØYSKOLE

Dette selvstendige arbeidet er gjennomført som ledd i masterstudiet i økonomi- og administrasjon ved Norges Handelshøyskole og godkjent som sådan. Godkjenningen innebærer ikke at Høyskolen eller sensorer innestår for de metoder som er anvendt, resultater som er fremkommet eller konklusjoner som er trukket i arbeidet.

Sammendrag

Formålet med denne utredningen er å undersøke bruk av karakterer i prestasjonsevalueringer av ansatte, hvor karakterene gis basert på skjønnsmessige vurderinger og forventes å følge en forhåndsbestemt karakterfordeling. Det er begrenset med litteratur som tar for seg slike prestasjonsevalueringer. Vi gjennomfører derfor en eksplorerende studie for å bidra med ny innsikt på området. For å belyse utredningens tema har vi en kvalitativ tilnærming med 15 dybdeintervjuer i tre operatørbedrifter i oljebransjen.

Studien avdekker både muligheter og utfordringer med bruk av skjønn i karaktergivning. Et hovedtrekk er at skjønnsmessige prestasjonsevalueringer er nødvendig for å se helheten i ansattes prestasjoner. Om ansatte blir sett av sin leder, får tilstrekkelig med tilbakemeldinger, samt har klagemulighet dersom de føler seg urettferdig behandlet, virker skjønnsmessige prestasjonsevalueringer godt egnet for å vurdere ansatte med komplekse arbeidsoppgaver.

Formålet med bruk av karakterer og tvungen karakterfordeling i prestasjonsevalueringer er å skape transparent link mellom ytelse og belønning, samt sikre likere ytelsesstandarder på tvers av bedriften. Imidlertid innebærer tvungen karakterfordeling flere utfordringer, avhengig av hvilken type som benyttes. Vi finner at et krav om gjennomsnittskarakter fører til at hovedvekten av ansatte mottar samme karakter. Dette skyldes at ledere spares for tid og ressurser, samtidig som de unngår konflikter som følge av å gi svake karakterer. Forhåndsbestemte prosentandeler for hvor mange ansatte som kan få hver karakter, kan resultere i karaktergivning basert på hva som passer i fordelingen, heller enn karakterer som reflekterer ansattes faktiske prestasjoner.

Bedrifter som benytter karakterer i prestasjonsevalueringer bruker ofte disse som grunnlag for å allokere belønninger. Tittelen i vår utredning er basert på Gneezy og Rustichini (2000) sin anerkjente studie, ”Pay enough or don’t pay at all”. Studien viser at svake finansielle insentiver kan svekke motivasjon og arbeidsinnsats, men at motivasjon og innsats øker når styrken i insentivene øker. Våre funn stemmer overens med dette, men vi avdekker samtidig at svake finansielle insentiver tilknyttet karakterene kan føre til at bedrifter mister sine flinkeste ansatte. Vi argumenterer dermed for at bruk av tid og ressurser på karaktergivning ikke har noen hensikt med mindre det differensieres i karakterer og ansatte belønnes deretter. Vår konklusjon er derfor ”Pay enough or don’t rate at all”.

Forord

Denne utredningen er gjennomført som en avsluttende del av det femårige masterstudiet i økonomi og administrasjon ved Norges Handelshøyskole (NHH). Utredningen er skrevet med utgangspunkt i våre hovedprofiler Økonomisk Styring og Strategi og Ledelse. Arbeidet ble gjennomført høsten 2016.

Prestasjonsevalueringer basert på skjønn og forhåndsbestemte karakterfordelinger har blitt mindre populært internasjonalt det siste tiåret. I Norge bruker derimot stadig flere bedrifter denne metoden for å evaluere ansattes prestasjoner. Temaet er hittil forsket lite på, og brorparten av eksisterende litteratur fokuserer på amerikanske bedrifter. Gjennom denne utredningen ønsker vi derfor å øke innsikten i hvordan slike systemer virker i en norsk kontekst. Vi håper vår utredning kan være verdifull for bedrifter som benytter, eller vurderer å benytte seg av, karakterer og forhåndsbestemte karakterfordelinger i prestasjonsevalueringer. Arbeidet med utredningen har vært spennende, lærerikt og utfordrende. Vi har lært å spille på hverandres styrker og har opplevd å utfylle hverandre godt.

Vi ønsker å rette en stor takk til bedriftene som har deltatt i studien og til hver enkelt respondent vi har intervjuet. De har vist stor samarbeidsvilje og bidratt til verdifull datainnsamling. Vi vil også takke vår veileder, Iver Bragelien, for et konstruktivt samarbeid og gode råd gjennom hele prosessen. Dette har utvilsomt bidratt positivt til utredningens kvalitet.

Bergen, januar 2017

Idunn Marie Svendsen

Birgitte Kielland

Innholdsfortegnelse

1. INNLEDNING 11

1.1 BAKGRUNN FOR STUDIEN 11
1.2 PROBLEMSTILLING 13
1.3 AVGRENSNINGER 13
1.4 UTREDNINGENS STRUKTUR 14

2. TEORETISK RAMMEVERK 15

2.1 PRINSIPAL-AGENT-TEORI 15
2.2 INSENTIVER OG BELØNNING 17
2.2.1 FASTLØNN 17
2.2.2 PRESTASJONSBASERT LØNN 17
2.2.3 IKKE-FINANSIELL BELØNNING 20
2.3 INDRE OG YTRE MOTIVASJON 21
2.4 PRESTASJONSEVALUERINGER 23
2.4.1 OBJEKTIVE PRESTASJONSMÅL 23
2.4.2 SUBJEKTIVE PRESTASJONSMÅL 25
2.4.3 RELATIV PRESTASJONSEVALUERING 30
2.4.4 TVUNGEN FORDELING 31
2.4.5 RETTFERDIGHET I PRESTASJONSEVALUERINGER 36
2.4.6 TILBAKEMELDINGER 37
2.5 OPPSUMMERING AV STUDIENS TEORETISK RAMMEVERK 38

3. METODE 39

3.1 FORSKNINGSDESIGN 39
3.1.1 FORSKNINGSTILNÆRMING 40
3.1.2 METODE FOR INNSAMLING AV DATA 41
3.1.3 FORSKNINGSSTRATEGI 41
3.2 INTERVJUENE 42
3.2.1 DATAINNSAMLING 42
3.2.2 VALG AV ENHETER OG INTERVJUOBJEKTER 44
3.2.3 INTERVJUGUIDE 47
3.2.4 INTERVJUPROSESSEN 48
3.3 ANALYSE AV DATA 49
3.3.1 KLASSIFISERING AV DATA 49
3.3.2 PRESENTASJON AV FUNN 50
3.4 EVALUERING AV METODE 51
3.4.1 VALIDITET 51
3.4.2 RELIABILITET 53
3.5 ETISKE HENSYN 55
3.6 OPPSUMMERING AV STUDIENS METODE 56

4. BRANSJEN OG CASENE 57

4.1 BRANSJEN OG CASENE 57
4.2 BEDRIFTENES SYSTEMER FOR BELØNNING OG PRESTASJONSEVALUERING 58
4.2.1 BEDRIFT A 58
4.2.2 BEDRIFT B 60
4.2.3 BEDRIFT C 64
4.2.4 OPPSUMMERING AV BEDRIFTENES SYSTEMER FOR BELØNNING OG PRESTASJONSEVALUERINGER 66

5. FUNN TIL STUDIENS PROBLEMSTILLING 67

5.1 FUNN I BEDRIFT A 67
5.1.1 SKJØNNSMESSIGE PRESTASJONSEVALUERINGER I BEDRIFT A 67
5.1.2 KARAKTERER OG TVUNGEN FORDELING I BEDRIFT A 72
5.1.3 OPPSUMMERING AV FUNN I BEDRIFT A 80
5.2 FUNN I BEDRIFT B 81
5.2.1 SKJØNNSMESSIGE PRESTASJONSEVALUERINGER I BEDRIFT B 81
5.2.2 KARAKTERER OG TVUNGEN FORDELING I BEDRIFT B 84
5.2.3 OPPSUMMERING AV FUNN I BEDRIFT B 90
5.3 FUNN I BEDRIFT C 91
5.3.1 SKJØNNSMESSIGE PRESTASJONSEVALUERINGER I BEDRIFT C 91
5.3.2 KARAKTERER OG TVUNGEN FORDELING I BEDRIFT C 94
5.3.3 OPPSUMMERING AV FUNN I BEDRIFT C 99

6. ANALYSE 101

6.1 SKJØNNSMESSIGE PRESTASJONSEVALUERINGER 102
6.2 TVUNGEN FORDELING 110

7. KONKLUSJON 125

7.1 SVAR PÅ PROBLEMSTILLINGEN 125
7.2. BEGRENSNINGER VED STUDIEN OG FORSLAG TIL VIDERE FORSKNING 129

8. LITTERATURLISTE 131

9. VEDLEGG 143

9.1 INTERVJUGUIDE 143
9.1.1 INFORMASJON SOM GIS I STARTEN AV INTERVJUET 143
9.1.2 SPØRSMÅL TIL LEDERE MED EVALUERINGSANSVAR 144
9.1.3 SPØRSMÅL TIL IKKE-LEDENDE ANSATTE UTEN EVALUERINGSANSVAR 146

Tabeller

TABELL 1 OPPSUMMERING AV STUDIENS TEORETISKE RAMMEVERK **38**

TABELL 2 INTERVJUOBJEKTER **46**

TABELL 3 OPPSUMMERING AV STUDIENS METODE **56**

TABELL 4 KARAKTERER PÅ ATFERD OG LEVERANSE **63**

TABELL 5 OPPSUMMERING AV BEDRIFTENES SYSTEMER **66**

TABELL 6 OPPSUMMERING AV FUNN I BEDRIFT A **80**

TABELL 7 OPPSUMMERING AV FUNN I BEDRIFT B **90**

TABELL 8 OPPSUMMERING AV FUNN I BEDRIFT C **99**

TABELL 9 OPPSUMMERING AV FORMÅL MED BEDRIFTENES SYSTEMER **101**

1. Innledning

1.1 Bakgrunn for studien

Medarbeiderne er sett på som bedrifters viktigste ressurs og en kilde til konkurransefortrinn (Barney & Wright, 1997; Brewster, 2000; Henderson, 2011; Pfeffer, 1998). Evaluering av ansattes prestasjoner, med hensikt å forbedre prestasjonene, er derfor en viktig del av bedrifters HR-systemer (Judge & Ferris, 1993; Murphy & Cleveland, 1995). Prestasjonsevalueringer omfatter evaluering av ansattes prestasjoner, kommunikasjon til ansatte angående hvordan de presterer i jobben og utarbeidelse av planer for forbedring (Aguinis, 2009). Hensikten med prestasjonsevalueringer er å gi tilbakemeldinger og å belønne ansatte (Lazear & Gibbs, 2015). Cleveland, Murphy og Williams (1989) påpeker at prestasjonsevalueringer brukes som grunnlag i en rekke administrative beslutninger. I denne studien fokuserer vi på hvordan bedrifter bruker prestasjonsevalueringer for å differensiere belønninger.

I de fleste bedrifter brukes skjønn i prestasjonsevalueringer (Murphy & Cleveland, 1995; Prendergast, 1999). Ved skjønnsmessige prestasjonsevalueringer avgjør lederens subjektive oppfatning vurderingen av de ansattes prestasjoner og atferd (Prendergast, 1999). Skjønnsmessige vurderinger er egnet for å gi et helhetlig bilde av ansattes prestasjoner, da leder kan ta hensyn til all relevant informasjon vedrørende ansattes jobbutførelse (Gibbs et al., 2004). Imidlertid har ledere kognitive skjevheter som ofte medfører systematiske feilvurderinger av ansattes prestasjoner. Tendensen til å skille for lite mellom sterke og svake ytere er særlig fremtredende. Dette er uheldig, da lav grad av differensiering i prestasjonsevalueringer fører til at relasjonen mellom innsats og belønning blir vagere (Prendergast, 1999).

Grunnet motvilje mot å differensiere når vurderinger baseres på skjønn, har forskere undersøkt former for prestasjonsevalueringer som kan øke differensieringen (Goffin et al., 1996). Enkelte studier indikerer at relative vurderinger fører til større skille mellom sterke og svake prestasjoner enn absolutte vurderinger (Heneman 1986; Nathan & Alexander, 1988; Wagner & Goffin, 1997). *Forced distribution rating systems,* heretter kalt *tvungen fordeling*, er en form for relativ prestasjonsmåling som er ment å sikre differensiering (Lazear & Gibbs, 2015). Tvungen fordeling kan praktiseres på ulike måter, men generelt innebærer det at

ansatte får karakterer som skal tilfredsstille forhåndsbestemte fordelinger (Stewart, Gruys & Storm, 2010). En metode er å gi ansatte karakterer basert på fastsatte prestasjonsstandarder, men kun tillate en gitt prosentandel av ansatte innenfor hver kategori (Stewart et al., 2010). En annen tilnærming er å kreve en gjennomsnittskarakter, der fordelingen rundt gjennomsnittet kan variere (Lazear & Gibbs, 2015).

General Electric, med leder Jack Welch, var en av de første som innførte tvungen fordeling av karakterer som evalueringsform, under navnet GE-modellen (Bates, 2003; Blume et al., 2009; Stewart et al., 2010). De benyttet et system med fastsatte prosentandeler som kunne havne innenfor hver kategori. Bedriften var lenge en forkjemper for tvungen fordeling, men gikk bort fra det etter flere tiår. Også Microsoft er kjent for bruken av tvungen fordeling, der de jevnlig rangerte ansatte i kategoriene topp, gode, middelmådige, mindre gode og dårlige medarbeidere (Eichenwald, 2012). Etter innføringen av tvungen fordeling opplevde Microsoft en rekke utilsiktede negative konsekvenser. Blant annet ble kollegaer mindre villige til å dele informasjon med hverandre, og flinke ansatte unngikk å jobbe med andre flinke i frykt for å få dårlig karakter (Eichenwald, 2012). Microsoft valgte på grunn av dette å gå bort fra tvungen fordeling, og legger nå større vekt på å fremme teamarbeid og medarbeidernes vekst og utvikling.

Generelt er det en nedadgående trend i bruken av tvungen fordeling internasjonalt (Jones & Rock, 2015). I tillegg til General Electric og Microsoft har mange store bedrifter valgt å gå over til mindre formaliserte evalueringsformer av sine ansatte (Jones & Rock, 2015). De Facto kunnskapssenter for fagorganiserte og HR Norge finner på en annen side at stadig flere bedrifter i Norge innfører karakterer og tvungen fordeling i prestasjonsevalueringer (Dahle, 2014). Det er derfor interessant å undersøke hvordan systemene virker i norske bedrifter.

I følge Stewart et al. (2010) er det lite forskning på tvungen fordeling både fra perspektivet til ledere som evaluerer og ikke-ledende ansatte. Videre stammer hovedvekten av studiene på området fra USA (Stewart et al., 2010). Det er derfor interessant å undersøke hvilke muligheter og utfordringer som oppstår når tvungen fordeling brukes i en annen kontekst. Dette gjelder særlig fordi USA og Norge skiller seg klart fra hverandre med tanke på bedriftskultur og HR-systemer (Hall & Soskice, 2001). I tillegg til manglende forskning på tvungen fordeling, er prestasjonslønn basert på skjønnsmessige vurderinger lite studert i litteraturen (Bragelien & Kvaløy, 2014). Dette er uheldig når halvparten av alle kompensasjonskontrakter baseres på skjønn (Bol, 2008). Fordi karakterer og

karakterfordelinger som regel fastsettes på bakgrunn av skjønnsmessige vurderinger (Blume et al., 2009), vil utredningen også ta for seg muligheter og utfordringer med skjønnsmessige prestasjonsevalueringer.

For å skaffe innsikt i hvordan ledere med evalueringsansvar og ikke-ledende ansatte oppfatter prestasjonsevalueringer basert på skjønn og tvungen fordeling av karakterer, gjennomfører vi, grunnet manglende empiri på området, en eksplorerende casestudie. Videre er det behov for en metode som går i dybden og får frem nyanser, for å sikre økt forståelse for ulike holdninger til systemene. Vi innhenter derfor data ved å gjennomføre semistrukturerte intervjuer av både ledere med evalueringsansvar og ikke-ledende ansatte.

1.2 Problemstilling

Formålet med studien er å skaffe innsikt i prestasjonsevalueringer basert på skjønn og tvungen karakterfordeling. Tre operatørbedrifter i oljebransjen, som alle har en arbeidsstokk bestående av kunnskapsarbeidere med høy grad av kompetanse og erfaring, utgjør studiens empiriske grunnlag. Alle bedriftene bruker skjønnsmessige vurderinger og tvungen fordeling av karakterer i evaluering av de ansatte. For å tilegne oss informasjon om hvordan slike prestasjonsevalueringer virker gjennomføres 15 dybdeintervjuer. Vi intervjuer både ledere med evalueringsansvar og ansatte uten evalueringsansvar, for å sikre et bredere perspektiv.

Vi har valgt følgende problemstilling for vår utredning:

Hvordan virker prestasjonsevalueringer basert på skjønn og tvungen fordeling av ansattes karakterer?

1.3 Avgrensninger

Denne utredningen undersøker bruk av skjønnsmessige prestasjonsevalueringer og tvungen fordeling av ansattes karakterer i tre operatørbedrifter i norsk olje- og gassektor. Studien er avgrenset til denne sektoren, som heretter benevnes oljebransjen. I utredningen gis det ikke konkrete forslag til hvordan systemer for prestasjonsevalueringer av ansatte bør utformes. Den begrenser seg til å undersøke virkninger av prestasjonsevalueringer basert på skjønn og tvungen karakterfordeling.

Prestasjonsevalueringer kan brukes i en rekke administrative beslutninger. I studien vil vi se på hvordan bedrifter bruker prestasjonsevalueringer som grunnlag for å allokere belønninger. Studien tar for seg både lederes holdning til å evaluere og gi ansatte karakterer, og ansattes holdning til å bli evaluert og motta karakterer som grunnlag for prestasjonsbasert lønn. Vi ser derimot bort fra lederavlønning og hvordan deres prestasjoner evalueres. Utredningen er videre avgrenset til å se på prestasjonsbasert lønn i form av lønnsøkninger og bonuser. Vi ser bort i fra andre former for finansielle-belønninger som pensjon og forsikringer.

1.4 Utredningens struktur

Denne utredningen er delt inn i syv kapitler. I kapittel 2 presenteres vårt teoretiske rammeverk. Hensikten er å danne en grunnforståelse for prestasjonsevalueringer basert på skjønn og tvungen fordeling av ansattes karakterer. I kapittel 3 beskrives metoden som er valgt for denne utredningen. Her argumenterer vi for ulike metodiske valg og hvordan disse gjør utredningen egnet til å besvare problemstillingen på best mulig måte. Videre blir det forklart hvordan vi har behandlet og analysert data, i tillegg til at utredningens validitet og reliabilitet blir drøftet.

Kapittel 4 gjør rede for kjennetegn ved bransjen som undersøkes, bedriftene som utgjør det empiriske grunnlaget og systemene de ulike bedriftene bruker for belønning og prestasjonsevalueringer. Det er relevant med slik innsikt, da dette danner rammebetingelsene for studien. I kapittel 5 presenteres empiriske funn fra hver av bedriftene. Basert på de empiriske funnene analyserer vi i kapittel 6 likheter og forskjeller mellom bedriftene opp mot vårt teoretiske rammeverk. Med bakgrunn i funnene fra analysen presenteres vår konklusjon i kapittel 7, der vi besvarer studiens problemstilling. Avslutningsvis presenteres forslag til videre forskning kombinert med studiens begrensninger.

2. Teoretisk rammeverk

I dette kapittelet presenteres utredningens teoretiske rammeverk. Vi vil beskrive og forklare teoretiske begreper og konsepter relatert til belønning og prestasjonsevaluering. Hensikten er å danne grunnforståelse for studiens tema. Først presenteres grunnleggende prinsipal-agent teori. Deretter introduseres teori og empiri relatert til belønning og insentiver. Dette vil skape en forståelse for motiv og konsekvenser ved bruk av ulike former for belønning. Videre presenteres teori og empiri tilknyttet indre og ytre motivasjon. Vi vil så ta for oss ulike former for prestasjonsevalueringer. Her trekkes det frem teori og empiri knyttet til objektiv, skjønnsmessig og relativ prestasjonsevaluering, samt tvungen karakterfordeling. Avslutningsvis presenteres betydningen av rettferdighet og tilbakemeldinger i prestasjonsevalueringer. Et bredt teoretisk og empirisk utgangspunkt styrker grunnlaget for å utforske skjønnsmessige prestasjonsevalueringer og tvungen karakterfordeling.

2.1 Prinsipal-agent-teori

I klassisk økonomisk teori antas mennesker primært å motiveres av egeninteresse, som i stor grad dreier seg om et ønske om økonomisk belønning. Denne tankegangen gir opphav til prinsipal-agent teori (Cappelen & Tungodden, 2012). Prinsipal-agent teori ble introdusert på 1970-tallet, og de mest kjente bidragsyterne er Jensen og Meckling (1976).

Jensen og Meckling (1976) definerer en prinsipal-agent relasjon som: ”*En kontrakt hvor en eller flere personer (prinsipalen) engasjerer en annen person (agenten) til å gjøre en jobb på deres vegne, og som innebærer at en form for beslutningsmyndighet er delegert til agenten*” *(s. 310).*

Et eksempel på en prinsipal-agent relasjon er forholdet mellom en arbeidsgiver og en arbeidstaker. Prinsipalens (arbeidsgivers) verdiskapning avhenger av innsatsen agenten (arbeidstaker) yter. Problemet oppstår når agenten fatter beslutninger på prinsipalens vegne, men prinsipalen ikke er i stand til å observere agentens handlinger perfekt og kostnadsfritt (Eisenhardt, 1989). Siden prinsipalen ikke har mulighet til å verifisere hvor mye innsats agenten yter, er det ikke mulig å skrive kontrakt direkte basert på innsatsnivå.

Det er her naturlig å trekke frem tre antagelser ved en prinsipal-agent relasjon; mulig overskudd, interessekonflikt og asymmetrisk informasjon (Hendrikse, 2003). Mulig overskudd må være tilstede om det skal være lønnsomt å inngå en relasjon. Prinsipalen må være villig til å betale agenten mer for å utføre oppgaven enn det vil koste agenten å utføre den. Videre vil det foreligge en interessekonflikt mellom prinsipalen og agenten. I prinsipal-agent teori antas begge parter å opptre ut fra egeninteresse, og derav forsøker de å maksimere egen nytte. Prinsipalen ønsker at agenten skal yte høyest mulig innsats, men agenten har innsatskostnader og ønsker dermed å yte minst mulig. En tredje antagelse er asymmetrisk informasjon mellom prinsipalen og agenten, da agenten har bedre innsikt i egne evner og prestasjoner. Motstridende interesser og asymmetrisk informasjon skaper i hovedsak to typer problemer: *moralsk hasard* og *ugunstig utvalg* (Hendrikse, 2003).

Moralsk hasard

Moralsk hasard oppstår som en konsekvens av asymmetri mellom prinsipalen og agenten. Prinsipalen vil ikke ha tilstrekkelig informasjon til å vite med sikkerhet hvilken innsats agenten har utøvd (Holmstrom, 1979). Prinsipalen ønsker at agenten skal yte høyest mulig innsats, men agenten ønsker ikke å bli påført ekstra innsatskostnader. Dette kan føre til at agenten velger handlinger som ikke er i prinsipalens interesse. Agenten må derfor gis insentiver til å yte høy innsats, hvor prestasjonsbasert lønn kan brukes som et virkemiddel for å få agenten til å handle i tråd med prinsipalens ønsker (Milgrom, 1992).

Ugunstig utvalg

Ugunstig utvalg oppstår når en agent har privat informasjon før et ansettelsesforhold. For prinsipalen vil den skjulte informasjonen medføre usikkerhet vedrørende agentens fremtidige innsats i bedriften. Uten tilstrekkelig informasjon vil ikke prinsipalen klare å skille høyproduktive potensielle arbeidstakere fra lavproduktive potensielle arbeidstakere (Greenwald, 1986). Konsekvensen kan bli at bedriften ender med å betale lik lønn til både høy- og lavproduktive ansatte. I så måte kan høyproduktive ansatte unngå å søke jobb i bedriften, da deres innsats ikke vil verdsettes tilstrekkelig. Dette gir grunnlag for å innføre prestasjonsbasert lønn, med hensikt å sortere ut ansatte man ønsker i bedriften (Langeland, 1999).

Kritikk av prinspipal-agent-teori

Prinsipal-agent teori er kritisert for et ensidig fokus på økonomiske belønninger. Arrow (1985) kritiserer teorien for å utelukke aspekter som sosial belønning og normer, som ved

siden av økonomiske belønning vil være viktige i prinsipal-agent relasjoner. Knyttet opp mot motivasjonsteori, mener Ryan & Deci (2000) at mennesker er moralsk motivert til å yte innsats selv i fravær av økonomisk belønning.

Prinsipal-agent teori tar utgangspunkt i ærlige prinsipaler som prøver å kontrollere atferden til agenter som ikke kan stoles på. Prendergast og Topel (1993) påpeker imidlertid at prinsipaler også kan opptre uærlig overfor sine agenter ved prestasjonsmåling, eksempelvis gjennom underrapportering av prestasjoner.

2.2 Insentiver og belønning

Valg av ulike insentiv- eller belønningsformer varierer med hensyn til hvordan de er ment å påvirke ansattes motivasjon, og med hvilken risikodeling de innebærer (Langeland, 1999). Belønninger kan både være finansielle og ikke-finansielle. Blant finansielle belønninger er fastlønn den vanligste, etterfulgt av variabel, prestasjonsbasert lønn, og en kombinasjon av disse. Finansielle belønninger kan forsterkes av ikke-finansielle belønninger (Armstrong, 2012). I det følgende presenteres ulike former for finansiell og ikke-finansiell belønning.

2.2.1 Fastlønn

Fastlønn kan defineres som et fastsatt pengebeløp som en ansatt mottar for å utføre et arbeid (Armstrong, 2012). Denne typen belønning betales uavhengig av prestasjon, og ligger i bunn for eventuelle variable lønnskomponenter (Nordhaug, 2002). Fastlønn blir gjerne satt på bakgrunn av utdanning, erfaring og arbeidsoppgaver, heller enn den enkeltes bidrag. Gjennom fastlønnen viser en bedrift hvilke ansatte den ønsker, og hva den er beredt til å betale de som er kvalifisert (Lazear & Gibbs, 2015).

2.2.2 Prestasjonsbasert lønn

Bruken av variable, prestasjonsbaserte lønnskomponenter har blitt mer vanlig i norsk næringsliv de siste 20 årene. Barth (2005) viser til at omtrent halvparten av norske, private bedrifter bruker en form for prestasjonsavlønning. Veksten i bruk av prestasjonslønn bekreftes av Lunde og Grini (2007), som gjennom sin undersøkelse finner at andelen norske heltidsansatte som fikk utbetalt bonus økte fra 18 til 28 prosent mellom 1997 og 2005.

Prestasjonsbasert lønn kan defineres som en form for belønning hvor: *".. hele eller deler av den økonomiske belønningen varierer med hvilke resultater de ansatte og/eller bedriften frembringer" (Colbjørnsen et al., 2000; s. 1).* Eksempler på prestasjonsbasert lønn kan være lønnsøkninger og bonuser, akkord, provisjon, aksjer og aksjeopsjoner. I henhold til Colbjørnsen et al. (2000) kan formål bak innføring av resultatavhengig lønn være å motivere til økt innsats, påvirke beslutninger og oppmuntre til å jobbe riktigere, stimulere til risikotaking, og påvirke rekruttering og sammensetningen av arbeidsstokken.

Utformingen av prestasjonsbasert lønn må tilpasses de formål en bedrift har med slike systemer (Colbjørnsen et al., 2000). En viktig avveining vil være hvor stor andel av den totale lønnen som skal være resultatavhengig. Det er sjeldent hensiktsmessig å gjøre hele lønnen variabel, da det eksponerer lønnsmottaker for stor økonomisk risiko. Spørsmålet vil derfor gjerne være hvor stor den variable lønnen skal være i forhold til fastlønnen. Videre **er** viktige avveininger hva slags resultater som skal ligge til grunn for lønnen, og om den skal være individuell eller gruppebasert (Colbjørnsen et al., 2000). Individuelle og gruppebaserte ordninger har sine styrker og svakheter, og er hensiktsmessige for ulike formål.

Individuell prestasjonsbasert lønn

Ved individuell, prestasjonsbasert lønn knyttes avlønningen til den enkeltes prestasjoner. Bragelien (2003) viser til at individuelle ordninger og ordninger for mindre team har størst insentiveffekt. Dette kan forklares av at man ser tydeligere sammenheng mellom egen prestasjon og belønning.

I henhold til Colbjørnsen et al. (2000) kan individuelle ordninger ha gunstig effekt på ansattes motivasjon i tilfeller hvor den enkeltes bidrag til resultatet er lett synlig, og lett å skille fra andre kollegaers bidrag. I tillegg kan individuelle ordninger ha motiverende effekt når den enkelte har myndighet til å fatte beslutninger som påvirker resultatene og målene. Individuell prestasjonslønn kan også benyttes som et verktøy for å tiltrekke og beholde ønsket arbeidskraft i organisasjonen (Werther & Davis, 1989). En sorteringseffekt kan oppstå, hvor talentfulle ansatte blir tiltrukket av jobben eller ønske å bli værende i jobben, mens mindre talentfulle ansatte kan få insentiver til å bytte jobb (Blakemore, Low & Ormiston, 1987; Lazear, 2000).

Jenkins, Gupta, Mitra og Shaw (1998) gjennomførte en studie av sammenhengen mellom insentiver og kvalitet og kvantitet i ytelse. Studien avdekket at individuell prestasjonsbasert belønning virker på kvantitet, men ikke på kvalitet, i utført arbeid. Det er uheldig i bedrifter hvor faktisk ønsket effekt av insentivsystemet er å øke kvaliteten på ansattes arbeid. Sammenhengen mellom økonomiske insentiver og kvantitet gjaldt i tillegg kun for enkle, rutinepregede og lett målbare oppgaver. Dette underbygges av en annen empirisk studie utført av Weibel, Rost og Osterloh (2007). De finner at prestasjonsbasert belønning øker prestasjoner i bedrifter hvor arbeidsoppgavene er enkle og uinteressante, men reduserer prestasjoner for komplekse og/eller interessante arbeidsoppgaver.

En annen ulempe med individuell prestasjonslønn er at det i mange tilfeller er vanskelig å måle innsatsen til den enkelte medarbeider (Holmstrom & Milgrom, 1991). I tillegg kan belønning av arbeidstakere for kun individuelle prestasjoner redusere samarbeid (Baker, 2000). I slike tilfeller kan gruppebaserte ordninger være å foretrekke fremfor individuelle ordninger.

Kollektiv prestasjonsbasert lønn

Ved kollektiv, prestasjonsbasert lønn knyttes avlønningen til gruppers eller organisasjonens samlede resultat, prestasjoner og/eller måloppnåelse. Gruppeinsentiver er gunstig når ansatte jobber tett sammen og det er vanskelig å måle individuelle bidrag separat (Alchian & Demsetz, 1972). I henhold til Armstrong (2012) kan gruppeinsentiver i slike situasjoner fremme samarbeid og kompetansedeling mellom ansatte i gruppen. Dette skyldes at når de som er avhengig av hverandre belønnes felles, har de insentiv til å jobbe sammen og dele kunnskap for å oppnå best mulig felles resultat (Armstrong, 2012). Bruk av gruppeinsentiver fremfor individuelle insentiver kan dermed medføre godt samarbeidsklima og bedre resultater.

Belønning basert på gruppers prestasjoner gir opphav til det økonomisk teori omtaler som *gratispassasjerproblemet* eller *1/N problemet* (Holmstrom, 1982). Når det er vanskelig å måle den enkeltes bidrag til måloppnåelsen har hvert gruppemedlem motiv til å være gratispassasjer på andres innsats. Den enkelte kan yte lav innsats uten at det påvirker egen belønning, fordi gruppen mottar belønning basert på felles resultater. Problemet gjelder spesielt for store grupper, da det vil være vanskeligere å identifisere ansatte med lav innsats enn i mindre grupper (Lazear, 1998).

Gratispassasjerproblemet kan imidlertid motvirkes av såkalt *peer pressure* (Kandel & Lazear, 1992). Gruppeinsentiver kan føre til at gruppens medlemmer motiveres til å følge med på hverandres prestasjoner og påvirke hverandre til å yte høyere innsats (Baron & Kreps, 1999). Dette gjelder særlig når kostnadene knyttet til overvåking av hverandres prestasjoner er lave (Kandel & Lazear, 1992). I henhold til Che og Yoo (2001) gjelder peer pressure spesielt ved gjentatt interaksjon mellom gruppemedlemmene, da medlemmer som ikke yter i henhold til forventningene kan straffes i senere perioder.

En studie utført av Blasi, Freeman, Mackin og Kruse (2008) viser at bedrifter med kollektive belønningsordninger i snitt har 3-5 prosent høyere økonomiske resultater enn virksomheter uten slike ordninger. I henhold til Bragelien (2012) kan imidlertid årsaken være at slike ordninger ikke innføres i bedrifter som går med underskudd, men heller i lønnsomme bedrifter. Videre påpeker Bragelien (2012) at observerte positive effekter av kollektive ordninger også kan skyldes eierskaps- og ansvarsfølelse overfor virksomhetens resultater.

2.2.3 Ikke-finansiell belønning

Latham og Locke (1979) påpeker at finansielle belønninger gjerne er det primære insentivet, men at det ikke alltid alene er nok til å motivere til høy innsats. Finansielle belønningsformer burde i henhold til Armstrong (2012) derfor støttes av ikke-finansiell belønning. Ikke-finansielle belønninger kan defineres som belønninger som ikke involverer direkte utbetalinger (Armstrong, 2012). Denne typen belønning oppstår ofte fra selve arbeidet, for eksempel gjennom prestasjon, anerkjennelse, autonomi, muligheter til å bruke og utvikle ferdigheter og karriereutvikling (Armstrong, 2007).

Anerkjennelse er ansett å være et av de kraftigste verktøyene for å belønne ansatte i en organisasjon (Armstrong, 2012). Denne typen belønning kan forekomme ved positiv og umiddelbar tilbakemelding fra ledere og kollegaer som anerkjenner individets eller teamets innsats. Anerkjennelse kan også vises ved at ledere hører på og handler etter forslag som kommer fra ansatte (Armstrong, 2012). Armstrong (2007) trekker også frem karriereutvikling som en viktig form for ikke-finansiell belønning for å motivere og belønne ansatte i organisasjoner. I sammenheng med karriereutvikling er det viktig å belønne gjennom å legge til rette for personlig utvikling.

2.3 Indre og ytre motivasjon

Motivasjon kan defineres som de biologiske, psykologiske og sosiale faktorene som forklarer hvorfor individer handler som de gjør (Kaufmann & Kaufmann, 2009). Bedrifter kan bruke motivasjon som et virkemiddel for å påvirke ansattes atferd i ønsket retning. For å oppnå ønsket atferd må imidlertid bedriften vite hva som motiverer ansatte i arbeidssammenheng (Buchanan, 2010). I klassisk organisasjonsteori har det blitt antatt at all atferd er motivert av ytre belønning (Skinner, 1953). Deci (1971) hevder imidlertid at stort fokus på ytre belønninger kan trekke oppmerksomheten vekk fra gleden ved å utføre et arbeid og mot resultater som følger av arbeidet.

I 1985 publiserte Deci og Ryan ”Self-Determination Theory”, hvor de gikk bort fra antakelsen om at all atferd er ytre motivert. De fokuserte i stedet på i hvilken grad individers atferd er selvbestemt og selvmotivert. I denne teorien skilles det mellom indre og ytre motivasjon, der skillet baseres på hvilken bakenforliggende årsak en handling har. Indre motivasjon refererer til atferd med bakgrunn i indre belønninger som interesse og tilfredshet, mens ytre motivasjon referer til atferd med bakgrunn i ytre belønninger (Ryan & Deci, 2000). I arbeidslivet tilsier indre motivasjon situasjoner hvor en ansatt utfører en handling uten ekstra påvirkning, mens ytre motivasjon vil si når den ansatte utfører handlingen med ønske om å oppnå ytre belønning som lønn og bonus.

Kognitiv evalueringsteori la grunnlaget for ”Self-Determination Theory, og gir en videre forståelse av begrepet indre motivasjon. Teorien tar utgangspunkt i de to grunnleggende psykologiske behovene kompetanseopplevelse og selvbestemmelse (Ryan & Deci, 2000). Kompetanse knytter seg til behovet for å lykkes i utførelsen av utfordrende oppgaver, derav et mestringsbehov. Selvbestemmelse eller autonomi er menneskers behov for å ta egne valg og selv være initiativtaker til handlinger. Ryan og Deci (2000) hevder disse behovene må tilfredsstilles for at mennesker skal være indre motiverte. Videre påpeker de at empiri har vist at tilhørighet også har positiv effekt på indre motivasjon.

Generelt har de fleste former for økonomisk belønning negativ effekt på indre motivasjon til å utføre oppgaver (Deci, Ryan & Koestner, 1999). Deci et al. (1999) fant i deres studie at studenter som mottok penger for å løse en oppgave rapporterte lavere grad av indre motivasjon til å gjennomføre oppgaven. Måten belønningen blir oppfattet på er imidlertid avgjørende for effekten på indre motivasjon. I tilfeller hvor ytre belønning oppfattes

kontrollerende kan belønningen svekke indre motivasjon (Deci et al., 1999). Deci et al. (1999) trekker imidlertid frem at dersom belønningsmottakeren oppfatter belønningen informerende om egen kompetanse, kan indre motivasjon styrkes. I henhold til Frey og Jegen (2001) kan i tillegg økonomiske belønninger som oppfattes støttende bidra til å styrke tilstedeværelsen av indre motivasjon, gjennom økt opplevelse av selvbestemmelse. Spontane belønninger kan også være mer hensiktsmessige enn forventede belønninger, da de ikke vil påvirke indre motivasjon negativt (Deci et al., 1999).

Gneezy og Rustichini (2000) gjennomførte en studie av sammenhengen mellom økonomiske insentiver og innsats. Studien avdekket at innføring av svake økonomiske insentiver kan føre til lavere arbeidsinnsats. Dette vil typisk være tilfelle for oppgaver som er preget av høy indre motivasjon, hvor et svakt økonomisk insentiv bidrar til å svekke den indre motivasjonen til å utføre oppgaven. På den andre siden viser studien at innsatsen øker når styrken i de økonomiske insentivene øker. Konklusjonen til Gneezy og Rustichini (2000) er derfor: *Pay enough or don´t pay at all*.

I henhold til Kuvaas (2005) avhenger hvilken type motivasjon som er den beste kilden til gode prestasjoner av egenskapene ved en arbeidsoppgave. Generelt er indre motivasjon mer effektivt for oppgaver som krever kvalitet, forståelse, utvikling og kreativitet (Gagné & Deci, 2005). Indre motivasjon er således viktig i kompetansesterke bedrifter og i bedrifter hvor tjeneste- og produktkvalitet er sentralt. Ytre motivasjon er på en annen side best egnet for trivielle, enkle og standardiserte oppgaver som i utgangspunktet ikke er indre motiverende (Jenkins et al., 1998; Weibel et al., 2010). For slike oppgaver kan bedrifter gjennom ytre belønningsformer betale for forutsigbar atferd. En studie utført av Cerasoli, Nicklin og Ford (2014) kom frem til samme konklusjon. Deres analyse viste at indre motivasjon er av størst betydning for kvalitative oppgaver, mens ytre belønninger er av størst betydning for kvantitative oppgaver.

2.4 Prestasjonsevalueringer

Til grunn for prestasjonsbaserte belønninger foreligger prestasjonsevalueringer. Prestasjonsevalueringer innebærer evaluering av ansattes prestasjoner, kommunikasjon til ansatte angående hvordan de presterer i jobben, og utarbeidelse av planer for forbedring (Aguinis, 2009). Hensikten med prestasjonsevalueringer er gjennom vurdering av kvantifiserbare eller ikke-kvantifiserbare prestasjonsmål å gi tilbakemeldinger og insentivere ansatte (Lazear & Gibbs, 2015). Prestasjonsevalueringer fullføres som regel i en årlig medarbeidersamtale mellom en ansatt og vedkommendes linjeleder, der periodens prestasjoner diskuteres samtidig som mål for neste periode fastsettes.

Effektive prestasjonsmål gir nøyaktige, informative og rettidige indikasjoner på individets bidrag til bedriftens verdiskapning, og innebærer lav risiko for den ansatte (Holmstrom, 1979). Et prestasjonsmål fungerer optimalt hvis det reflekterer en medarbeiders faktiske bidrag til bedriftens verdiskapning (Baker, Gibbons & Murphy, 1994). Ved ikke-kvantifiserbare måltall er det derfor avgjørende at leder har mulighet til å observere ansattes innsats og arbeidshverdag. Dersom prestasjonsmål benyttes i insentivkontrakter burde de være nært knyttet til bedriftens strategi for å sikre at ansattes atferd styres i riktig retning (Ittner, Larcker & Rajan, 1997). Dette omtales som strategisk kongruens, og skal sikre at prestasjonsmålingene tjener som rettleder for hva som er viktig for å sikre virksomhetens suksess (Kuvaas & Dysvik, 2012). Informasjonsverdien i et prestasjonsmål avhenger av støy og hvor godt det reflekterer den ansattes innsats.

Det er vanlig å skille mellom objektive og subjektive teknikker for prestasjonsmåling, der begge metodene har styrker og svakheter. Oppsummert hevder Kuvaas og Dysvik (2012) at valget mellom subjektive og objektive teknikker er en avveining mellom henholdsvis måling av uviktige indikatorer på en relativt nøyaktig måte, og evaluering av viktige indikatorer på en unøyaktig måte.

2.4.1 Objektive prestasjonsmål

Objektive prestasjonsmål er kvantifiserbare og ofte numeriske, som for eksempel salgstall eller antall produserte enheter. Det gis liten mulighet for diskusjon om målet er nådd eller ikke, og de kalles derfor ofte ikke-vurderende prestasjonsmål. Styrken ved slike

prestasjonsmål er at de enklere kan kobles til belønning, og i mindre grad preges av lederens subjektive oppfatninger (Lazear & Gibbs, 2015).

Vridningseffeker og måleproblemer

Objektive måltall blir imidlertid ofte kritisert for å være ufullstendige (Gibbs et al., 2004). I mange jobber er det ikke mulig å skrive kontrakt direkte på ansattes innsatsnivå eller bidrag til verdiskapningen. Det skyldes at verdiskapningen ofte innebærer dimensjoner det er vanskelig å måle eller telle. Eksempler på slike dimensjoner kan være samarbeidsevner, kvalitet på arbeidet, vilje til å hjelpe andre og opprettholdelse av kunde- og leverandørrelasjoner. En ansatt har typisk mange oppgaver, det som kalles multitasking, der noen vil være mer kvantifiserbare og enklere å måle enn andre (Holmstrom & Milgrom, 1991). Prestasjonsmål som ikke tar tilstrekkelig høyde for, eller ignorerer, enkelte dimensjoner ved ansattes arbeidshverdag kan gi skjeve insentiver (Gibbs et al., 2004). En fare er at objektive måltall ikke måler det som virkelig er verdiskapende eller det som uttrykker gode prestasjoner. I henhold til Murphy (2008) er det sjeldent et en-til-en forhold mellom objektive måltall og ansattes prestasjoner og atferd.

Når jobber er komplekse er kompensasjonskontrakter nesten alltid ufullstendige (Gibbs et al., 2004). I tilfeller der det kun er mulig å skrive kvantifiserbare mål til deler av den ansattes verdiskapning, og det er de kvantifiserbare målene som er tilknyttet prestasjonslønn, kan bedriften få utfordringer med vridningseffekter. Problemet er særlig fremtredende når den ansattes oppgaver er substitutter, det vil si at innsats på en oppgave ikke gjør den ansatte mer effektiv i den andre oppgaven. Den ansatte vil da typisk rette atferden mot oppgaven den blir belønnet for, og neglisjere andre ikke-målbare oppgaver som er viktige for verdiskapningen (Holmstrom & Milgrom, 1991; Murphy, 2008). Er oppgavene derimot komplementære kan man få innsats på begge oppgaver, selv om man kun belønner den ene oppgaven. Om bedriften skal bruke objektive prestasjonsmål burde de derfor designe jobber på en måte som gjør at ansatte har komplementære oppgaver om de skal fortsette å gi sterke insentiver (Holmstrom & Milgrom, 1991). Alternativt burde de dempe prestasjonsbasert belønning for å hindre multitaskingproblemer og skjeve prestasjonsmål.

På bakgrunn av potensielle vridningseffekter, mangler objektive måltall ofte kriterierelatert validitet, de fanger altså ikke opp det som er viktig for verdiskapning og gode prestasjoner (Murphy, 2008). Et eksempel på vridningseffekter kan være ledere som fokuserer på å bedre kortsiktig profitt på bekostning av langsiktige kunderelasjoner. Innføring av

prestasjonsbasert lønn gir således feil fokus om det som måles ikke reflekterer dimensjoner som er viktige for verdiskapningen. Ideelt sett burde derfor insentivkontrakter bruke all mulig informasjon om ansattes påvirkning på bedriftens verdiskapning og vekte den riktig, slik at insentiver er godt balansert mellom ulike dimensjoner av jobben (Baiman & Rajan, 1995; Bommer et al., 1995; Feltham & Xie, 1994; Holmstrom & Milgrom, 1991; Kaplan & Norton, 1996).

Manipulering

Kvantitative prestasjonsmål er også ofte utsatt for manipulering av ansatte, grunnet informasjonsasymmetri mellom ledere og ansatte (Baker, 1992; Baker et al., 1994; Holmstrom & Milgrom, 1991; Ittner, Larcker & Meyer, 2003). Ansatte vet ofte mer om omgivelsene deres enn lederne, og den informasjonen kan brukes til å manipulere prestasjonsmål til den sin fordel (Gibbs et al., 2004). Enkelte finner det derfor uegnet å kalle kvantifiserbare prestasjonsmål objektive.

Ukontrollerbare forhold og risiko

En annen svakhet med objektive prestasjonsmål er manglende hensyn til at ansattes prestasjoner kan påvirkes av ukontrollerbare forhold (Murphy & Cleveland, 1995). Ukontrollerbare forhold, eller støy, er hendelser utenfor den enkeltes kontroll som påvirker verdiskapningen. Om prestasjonen blir lavere grunnet støy, vil den ansattes innsats i mindre grad reflekteres av objektive måltall. Faste prestasjonsmål kan bli urimelig vanskelige å nå, noe som kan påvirke ansattes motivasjon negativt (Merchant & Manzoni, 1989). Særlig er utfordringen fremtredende når det er store finansielle eller ikke-finansielle konsekvenser ved å ikke nå målene, fordi ansatte da påføres stor nedsiderisiko (Gibbs et al., 2004). På den andre siden vil kvantitative prestasjonsmål heller ikke ta hensyn til støy som påvirker ansattes prestasjoner positivt, og man vil derfor i enkelte tilfeller premiere flaks heller enn innsats. Murphy (2008) hevder manglende hensyn til støy er en av de største feilene man kan gjøre i prestasjonsevalueringer.

2.4.2 Subjektive prestasjonsmål

En viktig karakteristika i de fleste insentivkontrakter er bruken av subjektivitet i evaluering og belønning av ansatte (Murphy & Cleveland, 1995; Prendergast, 1999). Subjektive prestasjonsmål er egnet til å gi et mer helhetlig bilde av ansattes prestasjoner, da leder kan ta hensyn til all relevant informasjon vedrørende ansattes prestasjoner (Gibbs et al., 2004). Med

subjektive vurderinger vil belønningen avgjøres av lederens skjønnsmessige oppfatning av ansattes prestasjoner og atferd (Prendergast, 1999). Skjønnsmessige prestasjonsmålinger gjør at lederen også kan evaluere aspekter ved ansatte som det er vanskeligere å kvantifisere, som samarbeidsevner, stå-på-vilje og initiativtaking (Hayes & Schaefer, 2000).

Skjønnsmessige prestasjonsevalueringer brukes ofte for å bøte på svakheter med objektive måltall og belønne for verdiskapende innsats det er vanskelig å kvantifisere. Bushman, Indjejikian og Smith (1996) argumenterer blant annet for at subjektivitet brukes for å redusere multitaskingproblemer og vridningseffekter som kan oppstå med objektive måltall. Når lederen bruker skjønn i prestasjonsevalueringen vil det ikke lenger lønne seg å neglisjere oppgaver som ikke er kvantifiserbare. Subjektivitet blir også brukt for å redusere sjansen for manipulering av kvantitative måltall (Murphy & Oyer, 2003).

Subjektivitet kan også brukes for å redusere ansattes risiko ved at støy kan filtreres ut (Gibbs et al., 2004; Murphy & Oyer, 2003). Det er en mer fleksibel tilnærming fordi insentivkontrakten og prestasjonsstandarder kan tilpasses etter hvert som omgivelsene endres (Merchant & Manzoni, 1989). Subjektive prestasjonsmål er således et godt alternativ for å holde ansatte motiverte når støy påvirker resultatene og faste prestasjonsmål blir urimelig vanskelige å nå. Gibbs et al. (2004) finner blant annet at subjektivitet brukes for å redusere ansattrisiko når de jobber i nedgangstider og har vanskelige mål. I henhold til tidligere studier brukes subjektivitet derimot sjeldent for å filtrere ut støy som påvirker ansattes resultater og lønn positivt (Merchant, 1989). En forklaring kan være at risikoaverse ansatte er mer opptatt av å bli forsikret mot nedsiderisiko (Gibbs et al., 2004). I tillegg kan ansatte oppleve filtrering av flaks som at ytelsesstandarder øker når prestasjonene er bedre enn forventet, noe som kan undergrave tillit (Gibbons, 1987). Subjektivitet brukes derfor oftere for å forsikre ansatte mot nedsiderisiko og filtrere ut effekter av uflaks heller enn flaks (Gibbs et al., 2004).

Ofte vil ansattes prestasjoner også påvirkes av støy som er kontrollerbart. Kontrollerbare forhold er hendelser som påvirker verdiskapningen, men i motsetning til ukontrollerbare forhold er sensitive til ansattes handlinger (Gibbs et al., 2004). Kontrollerbare forhold kan være tilfeldige, men ansatte kan likevel dempe negative effekter og utnytte positive effekter av støyet. Det er sjeldent entydig kontrollerbare eller ukontrollerbare forhold, de fleste kan delvis påvirkes av ansatte, og deres respons på støyet er derfor viktig for verdiskapningen.

Bruk av subjektivitet kan både redusere risiko ved ukontrollerbare forhold, men samtidig gi insentiv til å respondere så godt som mulig (Gibbs et al., 2004).

Utfordringer med subjektive prestasjonsmål

Utfordringen med bruk av skjønn i prestasjonsevalueringer er nettopp subjektiviteten som ligger i det. Ansatte utsettes her for en risiko for urettferdige prestasjonsevalueringer (Fulk, Brief & Barr, 1985; Prendergast & Topel, 1993). I henhold til Murphy (2008) er hovedproblemet med subjektive målinger at ledere verken har motivasjon eller kompetanse til å foreta nøyaktige vurderinger. Ledere ønsker ikke nødvendigvis å foreta unøyaktige vurderinger, men informasjonsmangel og kognitive skjevheter fører til ubevisste og systematiske feilvurderinger av ansattes prestasjoner (Murphy, 2008). Studier på området viser blant annet at egenskaper ved den som evaluerer, relasjonen mellom den som evaluerer og den som blir evaluert og andre egenskaper ved den som evalueres, har like stor betydning for evalueringen som selve prestasjonen (DeNisi & Pritchard, 2006; Levy & Williams, 2004).

Trynefaktor og uproduktiv påvirkning

Halo- og horneffekter i prestasjonsevalueringer refererer til situasjoner der lederen registrerer positive eller negative trekk ved en ansatt, og lar disse påvirke andre deler av vurderingen (Thorndike, 1920; Murphy et al., 1993). Saal et al. (1980) forklarer haloeffekter som tendensen til å evaluere helhetsinntrykket av en ansatt heller enn å skille mellom ulike prestasjonsdimensjoner.

Et eksempel på halo-og horneffekter er når trynefaktor påvirker prestasjonsevalueringer. Blant annet Prendergast og Topel (1993) påpeker at skjønnsmessige prestasjonsevalueringer kan være problematiske fordi leders evaluering påvirkes av relasjonen de har til sine ansatte. I følge tidligere forskning kan relasjonen mellom leder og ansatt ha like mye å si for prestasjonsevalueringen som selve prestasjonen (DeNisi & Pritchard, 2006; Levy & Williams, 2004; Spence & Keeping, 2010). I henhold til Sutton et al. (2013) kan opp til 60 prosent av variasjonen i prestasjonsevalueringer avgjøres av hvor godt leder liker den ansatte, altså av trynefaktor.

Særlig to konsekvenser har vist seg gjeldende når ledere favoriserer eller påvirkes av trynefaktor i sine evalueringer (Prendergast, 1999). For det første bruker agenter tid på å smiske som heller burde vært anvendt på produktive aktiviteter. For det andre kan

favorisering gi ineffektiv innsamling av informasjon om den enkeltes prestasjoner, slik at det blir vanskelig å avgjøre hvorvidt gode prestasjonsevalueringer stammer fra favorisering eller faktisk gode prestasjoner (Prendergast, 1999). I så måte vil det bli utfordrende når bedriften skal avgjøre hvem de skal forfremme og belønne.

Uproduktiv påvirkning har vært det primære fokuset i økonomisk litteratur vedrørende kognitive skjevheter i skjønnsmessige prestasjonsevalueringer. Relevant teoretisk arbeid på området inkluderer Holmstrom (1999), Milgrom (1988), Milgrom og Roberts (1988), Tirole (1992), og Allen og Gale (1992). Uproduktiv påvirkning refererer til aktiviteter ansatte gjør for å forbedre sine prestasjonsevalueringer, men som har mindre betydning for verdiskapningen enn andre aktiviteter (Prendergast, 1999). Ansattes smisking kan føre til at ledere overvurderer deres prestasjoner. Konsekvensen kan bli favorisering av ansatte som bruker mye tid på å smiske (Prendergast, 1999).

Ved utstrakt bruk av trynefaktor kan ansatte få mindre tillit til leders vurdering og oppleve den urettferdig. Risikoen for opplevd trynefaktor forsterkes i følge French (1968) av at mennesker har en tendens til å overvurdere egen betydning i en organisasjon. Også Salamin og Hom (2005) og Zenger (1992), hevder mennesker har en tendens til å vurdere egne prestasjoner som over gjennomsnittet. Ansatte som overvurderer egne prestasjoner kan oppleve misnøye om prestasjonsavlønningen blir lavere enn forventet, og oppleve trynefaktor fra leder selv om dette ikke er tilfelle.

Motvilje mot å differensiere

For å hindre misnøye blant ansatte kan leder unngå å skille tilstrekkelig mellom gode og dårlige prestasjoner (Prendergast, 1999). Teoretisk arbeid på skjønnsmessige prestasjonsevalueringer finner nettopp at ledere ikke differensierer tilstrekkelig mellom gode og dårlige prestasjoner (Prendergast & Topel, 1993). Det er særlig *ettergivenhets-bias* og *sentralbias* som er relevante når det gjelder sammenpressing av skalaen. Ettergivenhets-bias er når ledere overvurderer prestasjonen til svake ytere, og er motvillige mot å gi svake prestasjonsevalueringer. Sentralbias foreligger når ledere gir alle ansatte prestasjonsevalueringer rundt gjennomsnittet. Slike forvrenginger er godt dokumentert i teoretiske arbeid, der eksempelvis Landy og Farr (1980) og Murphy og Cleveland (1991) har dokumentert neglisjerbare differanser i evalueringer og kompensasjon mellom ansatte.

Mangel på tilstrekkelig differensiering knyttes til at ledere frykter svakere prestasjonsevalueringer kan gå på bekostning av relasjonen til sine ansatte, og medføre konflikter og misnøye (Bretz & Milkovich, 1989). Mangel på differensiering mellom sterke og svake prestasjoner gir feilaktig informasjon om prestasjoner (Guralnik et al., 2004; Jawahar & Williams, 1997). For både administrative og utviklende prestasjonsevalueringer kan mangel på differensiering være problematisk, da det gir organisasjoner svakere beslutningsgrunnlag ved eksempelvis forfremmelser, oppsigelser og identifisering av utviklingsbehov. I tillegg er det uheldig når ledere unngår å differensiere, fordi relasjonen mellom innsats og belønning blir vagere (Blume et al., 2009)

Litteraturen viser også at slik kompresjon av skalaen er mer utstrakt i situasjoner der evalueringen er viktig for avlønningen (Prendergast, 1999). Ledere finner det mer ubehagelig å gi dårlige evalueringer til ansatte hvis det medfører lønnsjusteringer, da evalueringen i så måte har større betydning for den ansatte (Prendergast, 1999). I tilfeller ledere avstår fra å differensiere preges de av bekvemmelighetshensyn. De evaluerer da på bakgrunn av hva som er enkelt og behagelig, heller enn å foreta informerte og grundige prestasjonsevalueringer. Dette reduserer sannsynligheten for at ansatte med utviklingsbehov identifiseres, og øker sjansen for at overvurderte ansatte blir forfremmet utover deres evner (Prendergast, 1999).

Underrapportering

Skjønnsmessige vurderinger kan ikke verifiseres av en tredjepart. Det kan medføre at prestasjonsmål blir forvrengt eller manipulert (Prendergast, 1999). En måte ledere kan forvrenge prestasjonsmålene på, er å underrapportere ansattes prestasjoner for å spare lønnsutgifter. Hvis lederen har rettmessig krav på bedriftens resultater vil enhver lønn utbetalt til agenten tas fra egen lomme. Ledere kan i slike tilfeller fristes til å underrapportere prestasjoner selv når de er svært gode, for å holde lønnsutgiftene nede (Prendergast, 1999).

Problemet med underrapportering kan reduseres av omdømmeeffekter. Dette fordi det på sikt vil være uheldig for formålet om å tiltrekke og beholde ønsket arbeidskraft hvis det blir kjent at ledere underrapporterer prestasjoner. I de fleste store bedrifter har imidlertid ikke ledere krav på bedriftens resultater, og det er derfor lave insentiver til å underrapportere prestasjoner for å senke kostnader (Prendergast, 1999).

Relasjoner og tillit

Bruken av subjektivitet er ikke alltid effektivt grunnet kognitive biaser og systematiske feilvurderinger (Gibbs et al., 2004). Fordi det involverer skjønn fungerer subjektivitet kun hvis lederen foretar rettferdige og informerte vurderinger, og hvis den ansatte godtar vurderingen. Dersom ansatte ikke stoler på at sine ledere foretar objektive og informerte vurderinger kan resultatet bli frustrasjon, svekket motivasjon og turnover (Gibbs et al., 2004). Problemene med subjektive prestasjonsmål reduseres derfor hvis leder og ansatt utvikler en god relasjon med stor grad av gjensidig tillit (Baker et al., 1994; McAllister, 1995; Murphy og Oyer, 2003; Pennings og Woiceshyn, 1987). Ved tilstrekkelig tillit mellom ledere og ansatte øker også belønningstilfredshet, og de positive sidene ved subjektive vurderinger vil overgå de negative (Gibbs et al., 2004). I henhold til Gibbs et al. (2004) avhenger grad av tillit blant annet av leders ansiennitet, hvor ansattes tilfredshet med prestasjonsevalueringer øker når lederen har sittet lenge. Høyere grad av tillit kombinert med større bruk av subjektivitet, øker ansattes tilfredshet med kompensasjonssystemer og gir større insentiver til å yte høy innsats (Gibbs et al., 2004).

2.4.3 Relativ prestasjonsevaluering

Grunnet motvilje mot å differensiere når vurderinger baseres på skjønn, har forskere undersøkt former for prestasjonsevalueringer som kan øke differensieringen (Goffin et al., 1996). Fokuset i denne forskningen har først og fremst vært på effektene av to generelle evalueringssystemer: absolutt og relativ prestasjonsevaluering. Ved absolutt prestasjonsevaluering blir individers prestasjoner vurdert mot en bestemt standard, mens i relative prestasjonsevalueringer blir individers prestasjoner sammenlignet med andres prestasjoner (Duffy & Webber, 1974). Enkelte studier indikerer at relativ prestasjonsevaluering er mer effektivt enn absolutt prestasjonsevaluering for å øke differensieringen i prestasjoner (Heneman, 1986; Nathan & Alexander, 1988; Wagner & Goffin, 1997). Bruk av relative vurderinger, hvor lønnspotten som skal fordeles ofte er fast, kan i tillegg føre til at skjønnsmessige vurderinger ikke misbrukes til å underrapportere prestasjoner for å spare lønnskostnader (Prendergast, 1999).

En styrke ved relative vurderinger er at det kan være enklere å observere relativ enn absolutt prestasjon (Lazear & Gibbs, 2015). I mange tilfeller vil det være enkelt å vurdere hvorvidt en ansatt presterer bedre eller dårligere enn andre, mens dens enkeltes bidrag til verdiskapningen kan være vanskelig å fastslå. I slike tilfeller kan relativ

prestasjonsevaluering være godt egnet som vurderingsform, da det kun er behov for informasjon knyttet til rangering av ansattes prestasjoner.

Relativ prestasjonsevaluering kan også brukes til å filtrere ut felles risikofaktorer i ansattes arbeidshverdag (Prendergast, 1999). Et eksempel kan være selgere med lignende jobber, der etterspørselen etter produktet de selger varierer av årsaker utenfor den enkeltes kontroll. Hvis ansatte kompenseres kun basert på absolutt prestasjon, vil de være eksponert for risiko knyttet til svingninger i etterspørselen. En løsning på problemet kan være å belønne ansatte ut i fra hvor godt de presterer relativt til andre som står overfor samme risikofaktorer. På denne måten vil man begrense effekten av tilfeldigheter og støy i målingen av ansattes prestasjoner (Lazear & Gibbs, 2015).

Ved at ansatte konkurrerer om en lønnspott ved relative vurderinger, vil den enkeltes lønn reduseres når andre gjør en god jobb. Dette kan gå på bekostning av godt samarbeid, da de ansatte får insentiver til å fremme seg selv, fremfor bedriften som helhet (Prendergast, 1999). I henhold til Lazear og Gibbs (2015) er dette en vesentlig utfordring med relative vurderinger, da man i de fleste jobber til en viss grad er avhengig av samarbeid med kollegaer. Når samarbeid er viktig bør derfor relative vurderinger unngås. Dette gjelder spesielt ved relative vurderinger i team. I slike tilfeller kan lik lønn til alle teamets medlemmer være å foretrekke, for å hindre sabotering av prestasjoner innad i teamet (Bose et al., 2010; Lazear, 1989).

2.4.4 Tvungen fordeling

Forced distribution rating systems, heretter kalt *tvungen fordeling*, er en form for relativ prestasjonsmåling som er ment å sikre differensiering (Lazear & Gibbs, 2015). Tvungen fordeling kan praktiseres på ulike måter, men generelt innebærer det at ansatte får karakterer som skal tilfredsstille forhåndsbestemte fordelinger (Stewart, Gruys & Storm, 2010).

Karakterer er en samlebetegnelse på summative evalueringer, som innebærer at vurderinger av ansattes prestasjoner oppsummeres i en tallkarakter eller verbal kategori (Kuvaas & Dysvik, 2012). Det vanlige er at ansatte evalueres både på atferd og leveranse (Dahle, 2014). Karakterer klassifiserer læringsnivå, og kan være nyttige til sorteringsformål. Bruk av karakterer begrunnes med at det for ansatte er viktig å vite på hvilket nivå de presterer. I tillegg kan karakterene virke motiverende, da ansatte stadig har noe spesifikt å jobbe mot (Dahle, 2014). Gjennom bruk av karakterer blir også de beste yterne synlige og enklere å

anerkjenne. Karakterene kan imidlertid svekke ansattes motivasjon hvis de oppleves som at lederne setter merkelapper på dem (Kuvaas & Dysvik, 2012). I henhold til Dahle (2014) virker verbale merkelapper tyngre for ansatte enn tallkarakterer.

Tvungen fordeling av karakterer kan utøves på ulike måter. Alle metodene innebærer at ledere vurderer ansattes prestasjoner med verbale kategorier eller tallkarakterer, og tilpasser karakterene for å tilfredsstille forhåndsbestemte krav. En metode for tvungen fordeling er å fastsette prosentandeler for hvor mange ansatte som kan få hver karakter (Stewart et al., 2010). Det typiske er at ansattes karakterer skal normalfordeles. Imidlertid skiller bedrifter seg fra hverandre når det gjelder nøyaktig hvilken prosentandel som forventes å motta hver karakter, og hvor mange karakterer som brukes. En annen tilnærming er å kreve en gjennomsnittskarakter hvor fordelingen rundt gjennomsnittet kan variere (Lazear & Gibbs, 2015). En tredje metode er å rangere ansatte ved å sammenligne hver enkelt sine prestasjoner med andre, og rangere ansatte fra beste til dårligste ytere (Stewart et al., 2010).

General Electric, med leder Jack Welch, var en av de første som innførte tvungen fordeling av karakterer som evalueringsform under navnet GE-modellen (Bates, 2003; Blume et al., 2009; Stewart et al., 2010). De benyttet et system med fastsatte prosentandeler som kunne havne innenfor hver kategori. Bedriften var lenge en forkjemper for tvungen fordeling, men gikk bort fra det etter flere tiår. Også Microsoft er kjent for bruken av tvungen fordeling, der de jevnlig rangerte ansatte i kategoriene topp, gode, middelmådige, mindre gode og dårlige medarbeidere (Eichenwald, 2012). Etter innføringen av tvungen fordeling opplevde Microsoft en rekke utilsiktede negative konsekvenser. Blant annet ble kollegaer mindre villige til å dele informasjon med hverandre, og flinke ansatte unngikk å jobbe med andre flinke i frykt for å få dårlig karakter (Eichenwald, 2012). Microsoft valgte på grunn av dette å gå bort fra tvungen fordeling, og legger nå større vekt på å fremme teamarbeid og medarbeidernes vekst og utvikling.

Generelt er det en nedadgående trend i bruken av tvungen fordeling internasjonalt (Jones & Rock, 2015). I tillegg til General Electric og Microsoft har mange store bedrifter valgt å gå over til mindre formaliserte evalueringsformer av sine ansatte (Jones & Rock, 2015). De Facto kunnskapssenter for fagorganiserte og HR Norge finner på en annen side at stadig flere bedrifter i Norge innfører karakterer og tvungen fordeling i prestasjonsevalueringer av de ansatte (Dahle, 2014). Karakterene knyttes gjerne til ansattes lønnsutvikling.

Fordeler med tvungen fordeling

Det er en rekke potensielle fordeler med å innføre tvungen fordeling av karakterer. For det første kan tvungen fordeling dempe feilvurderinger som kan oppstå når ledere fastsetter karakterer, som eksempelvis ettergivenhetsbias og sentralbias (Lazear & Gibbs, 2015; Stewart et al., 2010). Ettergivenhetsbias kan oppstå fordi ledere ikke ønsker å rettferdiggjøre karakterene til sine ansatte, ønsker å unngå konfrontasjon, bruker for lite tid på evalueringsprosessen, eller mangler opplæring i prestasjonsevalueringer (Stewart et al., 2010).

Tvungen fordeling ble utviklet i et forsøk på å ta tak i problemene med mangel på differensiering i prestasjonsevalueringer (McBriarty, 1988). Dette gjøres gjennom å tvinge ledere til å differensiere mellom sterke og svake ytere. Når tvungen fordeling utøves i form av krav om normalfordeling eller fastsatte prosentandeler som kan få hver karakter, kan ikke feilvurderinger som ettergivenhetsbias og sentralbias utbre seg. Tvungen fordeling, med fastsatte prosentandeler for antall ansatte som skal få hver karakter, kan derfor bidra til større spredning i karakterer (Bates, 2003; Blume et al., 2009; McBriarty, 1988; Stewart et al., 2010). Grunnet mangler på empiriske undersøkelser, er det derimot uklart hvordan virkningen på differensiering er når det forventes gjennomsnittskarakter heller enn fastsatte prosentandeler.

Tvungen fordeling kan også sikre at ansatte evalueres etter samme kriterier, slik at utfallet av evalueringsprosessen blir mer objektiv (Stewart et al., 2010). Ved bruk av karakterer og tvungen fordeling blir prosesser og ytelseskriterier mer formelt etablert, med belønning og avlønning direkte linket til ansattes karakter. Tvungen fordeling kan i tillegg tilrettelegge for mer åpen kommunikasjon mellom ledere og ansatte, slik at ansatte vet hvordan de ligger an og hva som må gjøres for å forbedre sine prestasjoner (Stewart et al., 2010). For ansatte er det svært uheldig å komme utfor en leder som ikke gir jevnlige og konstruktive tilbakemeldinger, da det blir vanskeligere å forstå hvorfor de havner i en eventuell ugunstig posisjon (Stewart et al., 2010).

Bruk av tvungen fordeling kan også bidra til administrative beslutninger, ved at bedrifter får identifisert sterke, normale og svake ytere (Stewart et al., 2010). Prestasjonsevalueringene kan på den måten brukes til å allokere prestasjonsbaserte belønninger og identifisere ansatte som burde forfremmes (Amalfe & Adelman, 2002). Dersom bedrifter ønsker å beholde ansatte som presterer godt, er differensiering mellom sterke og svake ytere nødvendig

(Stewart et al., 2010). MacDougall (1991) hevder turnover av bedrifters nøkkelpersoner og kritiske talenter minimeres når tvungen fordeling brukes til å identifisere og belønne de som driver bedriften fremover.

Et gjennomgående trekk i bedrifter er at ansatte som ikke presterer i henhold til ønskede standarder kontinuerlig flyttes rundt i bedriftene (Truby, 2001). I slike tilfeller kan tvungen fordeling hjelpe med å heve prestasjonsnivå ved å identifisere og eliminere ansatte med svake prestasjoner (Amalfe & Adelman, 2002; Truby, 2001). Dette er viktig fordi bedrifter med flere ansatte som presterer svakt kan trekke ned resten av bedriften (Vance & Davidhizar, 1998). En annen positiv side ved å identifisere svake ytere er at de ofte har feil stilling, heller enn at de presterer svakt på alle områder. Tvungen fordeling kan bidra til at dette kommer til overflaten (MacDougall, 1991), og sikre raskere overføring til andre stillinger. For de som presterer gjennomsnittlig kan utviklingsbehov identifiseres (Amalfe & Adelman, 2002).

Utfordringer med tvungen fordeling

Til tross for flere fordeler med tvungen fordeling, er det også utfordringer med slike systemer. Blant annet Jeffrey Pfeffer og Malcolm Gladwell hevder slike systemer kan ha uheldig virkning på bedriftskultur og prestasjoner (Pfeffer og Sutton, 2006). Samtidig har ledere gjennomgående mindre positive erfaringer med slike systemer enn andre former for prestasjonsevalueringer (Lawler, 2002). Schleicher, Bull og Green (2009) finner i tillegg at tvungen fordeling oppleves vanskeligere og mindre rettferdig enn andre evalueringsformer. Generelt har tvungen fordeling av ansattes karakterer vært lite brukt i Norge fordi det kan virke demotiverende og skape et usunt arbeidsmiljø basert på frykt og dårlig arbeidsmoral (Bates, 2003). Imidlertid har bruken av karakterfordelinger i norsk arbeidsliv økt de senere år (Dahle, 2014).

Som nevnt kan tvungen fordeling brukes for å sikre større grad av differensiering i prestasjonsevalueringer. Imidlertid finner Boyle (2001) og Stewart et al. (2010) det uheldig når systemet brukes som en unnskyldning for å gi ansatte svakere karakterer. Dette fordi tvungen fordeling kan gi ledere insentiv til å skylde på systemet heller enn å ta ansvar for å hjelpe svake ytere å forbedre seg. Ledere burde være gode nok til å ta de krevende samtalene, uten å ha et system som tvinger dem til å skille mellom sterke og svake prestasjoner (Boyle, 2001).

I tillegg er det sjeldent ansattes prestasjoner passer inn i en forhåndsbestemt fordeling, og karakterene reflekterer da ikke ansattes faktiske prestasjoner. Antakelsen om normalfordeling kan gjelde for store grupper med tilfeldige individer, men er sjeldent gjeldende for mindre grupper (Abelson, 2001). Et stort utvalg består av mellom 1000 og 1500 individer, men bedrifter bruker ofte normalfordeling av karakterer på grupper ned til 20-50 individer. Konsekvensen kan bli at ledere som har brukt opp plassene i enkelte kategorier tvinges til å gi ansatte med tilsvarende prestasjoner og resultater forskjellige karakterer (Stewart et al., 2010). Guralnik og Wardi (2003) hevder nettopp det at ansatte med like prestasjoner ender med ulike karakterer og store lønnsforskjeller er en kilde til urettferdighet ved tvungen fordeling.

Et eksempel på tilfeller ansatte med tilsvarende prestasjoner ender med ulike karakterer, kan være for ansatte som tilhører spesielt talentfulle grupper (Blume et al., 2009; Stewart et al., 2010). Slike ansatte straffes om en viss andel av dem må få svakere karakter enn de ville fått i en annen gruppe. Tvungen fordeling kan på den måten skape tapere selv i talentfulle grupper. En mulig effekt kan i følge Brown og Benson (2003) være at ansatte som får toppkarakterer ett år kan bli utmattet. Dette skyldes at forventningene som senere stilles for å gode karakterer vil være så høye at arbeidsmengden økes utover forsvarlige nivåer. Om funnene til Brown og Benson (2003) stemmer, kan tvungen fordeling ha en negativ effekt på ansatte som får de beste karakterene. Denne effekten er mer sannsynlig å inntreffe i tilfeller karakterene gis til små grupper enn i større grupper (Stewart et al., 2010). Ulempen er særlig diskutert vedrørende fastsatte prosentandeler av ansatte, men få har undersøkt om dette er gjeldende når det benyttes forventning om gjennomsnittskarakter.

Tilsvarende som det for ansatte i sterke grupper er ekstra krevende å få gode karakterer, vil ansatte i svakere grupper enklere oppnå bedre karakterer enn de fortjener. I slike grupper kan tvungen fordeling føre til at ansatte får mindre insentiv til å strekke seg, om de allerede får toppkarakter for gjennomsnittlige eller svake prestasjoner (Boyle, 2001). Dette er uheldig når bedrifter ofte implementerer tvungen fordeling nettopp for å heve ytelsesstandarder. Mange bedrifter bruker tvungen fordeling for å sikre en ønsket prosentandel som skal falle innenfor hver kategori, men det er høyst usikkert om de ender med riktige ansatte i de ulike kategoriene (Stewart et al., 2010).

Ovennevnte utfordring gjelder ikke kun for grupper innad i bedrifter, men også for bedrifter som helhet (Olson & Davis, 2003). Et eksempel kan være bedrifter som kun ansetter svært

dyktige og erfarne medarbeidere. Hvis tvungen fordeling benyttes i slike bedrifter, må en viss andel få svake karakterer til tross for gode eller gjennomsnittlige prestasjoner. Karakterene kan da oppleves tilfeldig gitt for å tilfredsstille fordelingen heller enn å reflektere faktiske prestasjoner (Olson & Davis, 2003). Schrage (2000) hevder blant annet at det blir mindre og mindre nyttig å bruke tvungen fordeling og differensiere ansattes belønninger jo mer man utvikler ansatte i bedriften.

I henhold til Pfeffer og Sutton (2006) burde bedrifter gi prestisje og status til mange heller enn få, og unngå store gap mellom de som får mest og de som får minst om man ønsker at medarbeiderne skal samarbeide. McBriarty (1988) finner blant annet at en ulempe med tvungen fordeling er at det kan bidra til usunn konkurranse mellom ansatte. McBriarty (1988) og Truby (2001) hevder tvungen fordeling undergraver teamarbeid, særlig når fordelingen brukes til å allokere belønninger.

I henhold til Pfeffer og Sutton (2006) er det ikke bevist at bedrifters prestasjoner og resultater blir bedre med bruk av tvungen karakterfordeling. De finner det lite hensiktsmessig å definere noen få som toppytere og resten som mindre gode. Tvungen fordeling kan eksempelvis medføre at selv sterke ytere unngår å ta på seg oppgaver som innebærer stor risiko, for å unngå svakere karakterer. Dette kan føre til mangel på utvikling og hindre bedriftssuksess (Guralnik, Rozmarin & So, 2004). Schrage (2000) hevder det er svært uheldig at tvungen fordeling ikke bare kan bidra til å demotivere gjennomsnittlige ansatte, men også de beste yterne.

2.4.5 Rettferdighet i prestasjonsevalueringer

Manglende hensyn til hva som oppleves rettferdig er en fallgruve i prestasjonsevalueringer. Når prestasjonsmålinger brukes til å avgjøre fordeling av ressurser og belønning er det avgjørende at ansatte opplever distributiv rettferdighet (Adams, 1965; Colquitt et al., 2001). Distributiv rettferdighet omtales ofte også som utfallsrettferdighet, og refererer til hvorvidt fordelingen av ressurser og belønning oppleves rettferdig.

I tillegg er det viktig med prosedyrerettferdighet, om ulikheter i fordelinger skal oppleves rettferdig. Prosedyrerettferdighet defineres som rettferdighet i prosessene som fører til allokeringen (Greenberg, 1990). Slik rettferdighet kan styrkes ved å ha holdbare og pålitelige målinger av prestasjoner, ansattes involvering i prosessene, og at systemet oppleves legitimt (Leventhal, 1980). Opplevelsen av prosedyrerettferdighet påvirkes også av nøyaktighet i

målingene, at evalueringene ikke preges av biases og hvorvidt ansatte vises respekt når tilbakemeldingene gis, uavhengig av prestasjonsnivå (Leventhal, 1980). Cawley et al. (1998) poengterer viktigheten av rettferdighet ved å hevde at selv de mest tekniske, sofistikerte og nøyaktige prestasjonsevalueringssystemene blir ineffektive om de ikke aksepteres og støttes av de ansatte.

Klagemulighet

Innføring av ordninger for prestasjonsbasert belønning og prestasjonsevaluering er krevende, fordi det alltid er noen som vil føle seg urettferdig behandlet. Dette forsterkes når mennesker har en tendens til å overvurdere egen betydning i organisasjonen (French, 1968). Bragelien (2005) hevder en viss misnøye ved prestasjonsbasert belønning er uunngåelig. Et tiltak som likevel kan redusere misnøyen og øke opplevelsen av rettferdighet, er å innføre klagemulighet på evalueringen og den prestasjonsbaserte belønningen. Muligheten til å klage kan dempe problemene fordi ansatte selv tar stilling til om de godtar utfallet eller ikke (Bragelien, 2005) Klagemulighet fungerer derfor som en sikkerhetsventil mot urettferdige prestasjonsevalueringer med tilknyttet prestasjonsbasert belønning.

2.4.6 Tilbakemeldinger

Maddux (1997) viser at over halvparten av kontoransatte ikke forstår hvordan deres arbeid vurderes. Tilbakemeldinger på prestasjoner er derfor vesentlig om ansattes atferd skal styres av prestasjonsevalueringer. Ledere og ansatte har ofte ulike oppfatninger av prestasjoner. Uten samtaler mellom ledere og ansatte om prestasjonene vil de ikke kunne korrigere hverandres oppfatninger og skape lik forståelse (Locke & Latham, 2002)).

Ved prestasjonsevalueringer er det vanlig å vurdere ansatte opp mot fastsatte mål. For at målene skal fungere effektivt, er det sentralt med tilbakemeldinger på hvordan man presterer i forhold til målene (Locke & Latham, 2002). Ved tilbakemeldinger får ansatte mulighet til å justere atferden for å nå målene på en bedre måte. Dette er særlig viktig når vurderingene er skjønnsmessige, da de ofte er vanskeligere å synliggjøre for ansatte enn når vurderingene er kvantitative. Konkrete og tydelige tilbakemeldinger er derfor nødvendig om målsetting skal gi uttelling på prestasjoner (Locke & Latham, 2002). Tilbakemeldinger og mål er gjensidig avhengig av hverandre, og virker bedre sammen enn hver for seg.

I henhold til DeNisi & Kluger (2000) er tilbakemeldinger avgjørende både for motivasjon og prestasjoner. Dette støttes av Erez (1977) og Becker (1978) som i sine studier fant høyere prestasjoner i grupper som mottok tilbakemeldinger enn i grupper som ikke fikk tilbakemeldinger. Det er imidlertid enklere for ledere å gi ansatte ros når resultatene er gode enn å gi god og konstruktiv kritikk på resultater under forventning. Bekvemmelighet kan gjøre ledere motvillige til å gi kritikk, da de ønsker å unngå konflikter. I så måte gir ledere ofte bedre tilbakemeldinger enn fortjent for å unngå å sette seg i en vanskelig situasjon. Dette er ikke ubegrunnet, da blant annet Aguinis et al. (2012) har vist at negative tilbakemeldinger kan medføre defensive reaksjoner og konflikter. En annen risiko med negative tilbakemeldinger er at de ofte oppleves mindre treffsikre og rettferdige (Aguinis et al., 2012)

2.5 Oppsummering av studiens teoretisk rammeverk

I dette kapittelet har vi presentert utredningens teoretiske rammeverk. Relevant litteratur omfatter prinsipal-agent-teori, teori om insentiver og belønning, motivasjonsteori og ulike former for prestasjonsevalueringer av ansatte. Et bredt teoretisk og empirisk utgangspunkt har styrket vårt grunnlag for å utforske skjønnsmessige prestasjonsevalueringer og tvungen karakterfordeling. Tabellen oppsummerer gjennomgått teori og empiri.

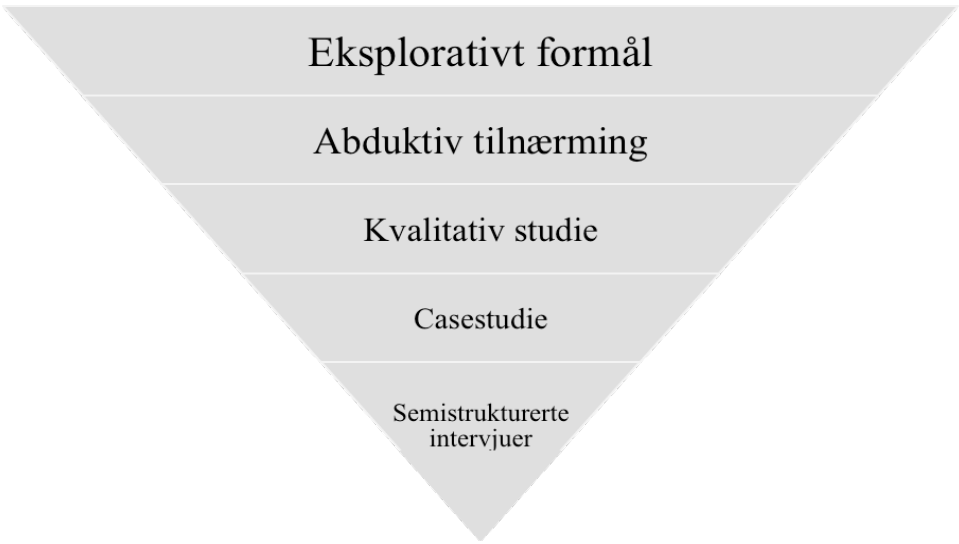
Teoretisk rammeverk	
Prinsipal-agent-teori	Moralsk hasard Ugunstig utvalg
Insentiver og belønning	Fastlønn Prestasjonsbasert lønn Ikke-finansielle belønninger
Motivasjon	Indre og ytre motivasjon
Prestasjonsevaluering	Objektive prestasjonsmål Subjektive prestasjonsmål Relativ prestasjonsevaluering Tvungen fordeling Rettferdighet i prestasjonsevalueringer Tilbakemeldinger

Tabell 1: Oppsummering av gjennomgått teori

3. Metode

I dette kapittelet beskriver og argumenterer vi for valg av metodisk tilnærming til studien. Den metodiske tilnærmingen beskriver teknikker, verktøy og prosedyrer som anvendes for å samle inn og analysere data (Saunders, Lewis & Thornhill, 2012). Valg av metode er gjort med hensikt å besvare vår problemstilling om skjønnsmessige prestasjonsevalueringer og tvungen fordeling.

Først presenteres valg av forskningsdesign og forskningstilnærming, som videre legger grunnlag for datainnsamling og -analyse. Til slutt evalueres datagrunnlagets validitet og reliabilitet. Figuren oppsummerer valg av metode i studien.



Figur 1: Metodisk tilnærming

3.1 Forskningsdesign

Forskningsdesign er en generell plan for hvordan man skal gå frem for å samle inn data og besvare problemstillingen. Forskningsdesignet skal inneholde klare mål, grunnlag for datainnsamling, begrensninger ved studien, og en diskusjon om etiske hensyn (Saunders et al., 2012). Valg av forskningsdesign bør ta utgangspunkt i formålet med studien, samt hvor mye forskning som eksisterer på området.

Det skilles i hovedsak mellom eksplorerende, beskrivende og forklarende forskningsdesign. Formålet med beskrivende studier er å skildre en nøyaktig beskrivelse av bestemte objekter, situasjoner eller prosesser (Jacobsen, 2015). Forklarende studier søker derimot å avdekke

kausale årsakssammenhenger (Saunders et al., 2012). Hensikten med eksplorative studier er å få en dypere forståelse for et fenomen, og slike studier er spesielt nyttig hvis man ønsker å oppnå klarhet og mer kunnskap (Jacobsen, 2015).

Hensikten med vår studie er å undersøke prestasjonsevalueringer basert på skjønn og tvungen fordeling av ansattes karakterer. Bruk av skjønn i evalueringsprosesser er lite studert i litteraturen (Bragelien & Kvaløy, 2014). I tillegg er det behov for flere studier om tvungen karakterfordeling, da det kun er utført et fåtall empiriske studier på området, og de få studiene som er gjennomført er ofte fra USA (Blume et al., 2009; Schleicher et al., 2009; Stewart et al., 2010). Videre ser vi mangel på forskning vedrørende ulike typer tvungen fordeling, der de fleste studier ser på systemer hvor en gitt prosentandel av ansatte må falle innenfor hver kategori. Denne studien vil også ta for seg systemer som krever en gjennomsnittskarakter heller enn fastsatte prosentandeler. Samlet virker et eksplorerende forskningsdesign hensiktsmessig fordi det gir oss muligheten til å erverve ny innsikt på områder med begrenset forskning, samt gi dypere forståelse for studiens tema. En styrke med forskningsdesignet er at det tilrettelegger for en fleksibel forskningsprosess, som gjør det mulig å revidere problemstillingen underveis i studien, etter hvert som funn av særskilt interesse avdekkes (Saunders et al., 2012).

3.1.1 Forskningstilnærming

Saunders et al. (2012) viser til to ulike forskningstilnærminger, henholdsvis *deduktiv* og *induktiv* tilnærming. Hvor klart man ser for seg teorien i starten av forskningen, reiser et viktig spørsmål om hvilken tilnærming som er mest hensiktsmessig (Saunders et al., 2012). En deduktiv tilnærming innebærer at man går fra det generelle til det spesielle. Teori legges i slike tilfeller til grunn for å utvikle hypoteser, som man senere tester ved analyse av data. En induktiv tilnærming innebærer at man går fra det spesielle til det generelle. I slike tilfeller kan teori og hypoteser formuleres på bakgrunn av observasjoner og empiriske funn (Saunders et al., 2012).

I denne utredningen er en induktiv tilnærming egnet, da vi tar sikte på å formulere ny teori og hypoteser på bakgrunn av empiriske funn (Saunders, et al., 2012). Samtidig vil det være nyttig å ha en teoretisk forankring i eksisterende teori, noe som taler for en deduktiv tilnærming. Det vil derfor være hensiktsmessig å benytte en kombinasjon av induksjon og deduksjon, det vil si en abduktiv tilnærming. Bakgrunnen er at vi ønsker å utforske hvorvidt

våre funn samsvarer med eksisterende teori, samtidig som vi vil se etter nye aspekter og sammenhenger. En abduktiv tilnærming gir oss muligheten til å tilføre ny innsikt i allerede kjente fenomener (Aase & Fossåskaret, 2007), og egner seg derfor til å studere skjønnsmessige prestasjonsevalueringer og tvungen fordeling.

3.1.2 Metode for innsamling av data

I datainnsamlingsprosessen er det viktig å ta stilling til hvilken type data man vil innhente, og hvordan man skal samle inn informasjon. Saunders et al. (2012) beskriver forskningsmetoder som kvantitative eller kvalitative. Kvantitativ metode vil si datainnhenting eller dataanalyse som genererer eller bruker numeriske data. Kvalitativ metode er derimot datainnhenting eller dataanalyse som baserer seg på ikke-numeriske data. Metoden muliggjør dyptgående innsikt i kompliserte sammenhenger.

Utformingen av problemstillingen og valg av forskningsdesign legger føringer for valg av forskningsmetode. Vårt studie vil gjennomføres ved å benytte en kvalitativ tilnærming. En kvalitativ metode er særlig egnet til å avdekke sammenhenger i innsamlet data for å øke forståelsen for et fenomen (Saunders et al., 2012). I forbindelse med studiens eksplorerende design er det behov for en metode som går i dybden, får frem nyanser ved dataene som samles inn og er åpen for kontekstuelle forhold (Jacobsen, 2015). Når responsen kommer i form av ord, vil menneskelige erfaringer og resonnement komme tydeligere frem og kunne bidra til å gi et helhetlig bilde av skjønnsmessige prestasjonsevalueringer og tvungen fordeling. I henhold til Saunders et al. (2012) egner en kvalitativ tilnærming seg best når man undersøker få enheter i dybden. En kvalitativ metode er derfor forenlig med vårt eksplorerende design og abduktive tilnærming.

3.1.3 Forskningsstrategi

Forskningsstrategi er en plan for hvordan man skal besvare studiens problemstilling. En casestudie kan defineres som *"en forskningsstrategi som involverer en empirisk undersøkelse av et bestemt midlertidig fenomen innenfor dens virkelige kontekst, ved bruk av multiple kilder"* (Saunders et al., 2012, s. 666). Denne forskningsstrategien er hensiktsmessig når man søker dybdeinnsikt på et område, og legger til rette for nye og uventede funn. På bakgrunn av studiens eksplorerende design synes casestudie derfor godt egnet. Et annet argument for å benytte en slik forskningsstrategi er at det er en fleksibel tilnærming som kan revideres etter hvert som funn avdekkes (Ghauri & Grønhaug, 2010). En

ulempe med casestudier er at funn ikke kan generaliseres på grunn av et begrenset utvalg (Saunders et al., 2012). Imidlertid er ikke hensikten med studien å generalisere, men gi økt forståelse for prestasjonsevalueringer basert på skjønn og tvungen karakterfordeling.

Vi vil gjennomføre en multippel casestudie av tre bedrifter i olje- og gassbransjen. Dette gir oss mulighet til å undersøke problemstillingen i den aktuelle konteksten og hente inn data fra flere kilder for å oppnå en bredere forståelse (Yin, 2009). Samtidig gjør en multippel casestudie oss i stand til å undersøke likheter og forskjeller i de ulike casene.

3.2 Intervjuene

Dette delkapittelet tar for seg datainnsamling, valg av intervjuobjekter, presentasjon av intervjuguiden, samt inngående informasjon om intervjuprosessen.

3.2.1 Datainnsamling

Vi baserer vår kvalitative forskning på primærdata i form av 15 dybdeintervjuer, der hvert intervju varte omtrent 1,5 timer og ble tatt opp på bånd. Intervjuer anses egnet da vi får mulighet til å avdekke det unike ved hver respondent, samt innblikk i refleksjoner og opplevelser rundt skjønnsmessige prestasjonsevalueringer og tvungen fordeling. En fordel med slike intervjuer er at det gir god mulighet til å gå i dybden på temaet, samt gi forskerne en utvidet forståelse for konteksten av studien (Saunders et al., 2012). Videre kan intervjuer gi en nyanserikdom som kan gi grunnlag for å indikere årsakssammenhenger og belyse nye aspekter, noe som er i tråd med studiens eksplorerende form. Yin (2009) hevder intervjuer er en av de viktigste kildene til informasjon i en casestudie.

Intervjuer er også nyttige for å få personlig kontakt med intervjuobjektene (Saunders et al, 2012). Vi kan opplyse om hvordan informasjonen blir brukt og at både intervjuobjekter og bedriftene vil behandles anonymt. Det gir respondentene mer tillit til studien og gjør dem dermed mer villig til å dele tanker og meninger om sensitive temaer (Saunders et al., 2012). Dette er særlig viktig i vårt forskningsprosjekt, da personlige refleksjoner og opplevelser av bedriftenes belønningssystemer og prestasjonsevalueringer vil være sensitivt. Tilgang til data var dermed et viktig argument ved valg av innsamlingsmetode.

I henhold til Saunders et al. (2012) er intervjuer enten strukturerte, semi-strukturerte eller ustrukturerte. Ved strukturerte intervjuer baserer intervjueren seg på forhåndsbestemte og

standardiserte spørsmål, der svaralternativene gjerne er definert på forhånd (Saunders et al., 2012). En fullstendig pre-strukturering kan medføre at man mister viktige nyanser, da det kan være en form for lukking av data (Jacobsen, 2015). Ustrukturerte intervjuer vil være mer uformelle og uten forhåndsdefinerte spørsmål, der målet er en dyp forståelse for temaet man undersøker. Her er det viktig å ha en klar oppfatning av hvilke aspekter man ønsker å undersøke (Saunders et al., 2012). Ustrukturerte intervjuer kan imidlertid bli svært komplekse og krevende, og i noen tilfeller umulige, å analysere (Jacobsen, 2015). Vi har valgt å basere oss på semistrukturerte intervjuer da det er en mellomting mellom strukturerte og ustrukturerte intervjuer, og derfor egnet til å møte utfordringer ved de to ovennevnte.

På bakgrunn av presentert teori og empiri utformet vi en intervjuguide med temaer og spørsmål vi ønsket å avdekke. Intervjuguiden presenteres nærmere i avsnitt 3.2.3. Den ble dannet med hensikt å sikre fremdrift og struktur i intervjuene, samt sikre innsikt på de mest vesentlige områdene. Med semistrukturerte intervjuer hadde vi også mulighet til å fravike fra intervjuguiden og la intervjuobjektene snakke fritt i henhold til egne tanker, stille oppfølgingsspørsmål og eventuelt fjerne spørsmål om det var nødvendig (Saunders et al., 2012). Dette var sentralt da forskningens eksplorerende design gjorde at vi måtte være i stand til å håndtere uventede forhold som fremkom underveis i prosessen, og utforske relevante områder for vår studie. Fleksibiliteten sikret at vi ikke mistet informasjon som kunne bidra til å forstå viktige dimensjoner ved skjønnsmessige prestasjonsevalueringer og tvungen karakterfordeling. Semistrukturerte intervjuer er derfor godt egnet til vår kvalitative forskning (Jacobsen, 2015).

Semistrukturerte intervjuer krever at man i forkant av intervjuene setter seg godt inn i konteksten, bedriftene og relevant teori. Videre vil prosessen under og i etterkant av intervjuene være krevende, særlig gjennom transkribering og analyse. Det vil derfor være en tidkrevende prosess å benytte en slik datainnsamlingsmetode. Studiens validitet og reliabilitet vil også kunne diskuteres ved bruk av semistrukturerte intervjuer, noe vi kommer tilbake til i kapittel 3.4. Imidlertid har vi opplevd at studiens design har gjort oss i stand til å få god kjennskap til datamaterialet og en helhetlig forståelse for temaet vi forsker på. På bakgrunn av dette mener vi at innsamlingsmetoden har vært fordelaktig for å finne og analysere funn.

3.2.2 Valg av enheter og intervjuobjekter

I det følgende vil studiens empiriske grunnlag gjøres rede for. Dette innebærer hvordan vi har fått tilgang til data, valg av studieobjekter og valg av intervjuobjekter.

Tilgang til data og forarbeid

Etter vi hadde besluttet temaet vi ønsket å studere, startet vi med å finne ut hvilke bedrifter i bransjen som benytter skjønnsmessige prestasjonsevalueringer og tvungen karakterfordeling. Til dette formål utførte vi tre bakgrunnsintervjuer med ulike personer i oljebransjen. Gjennom intervjuene fikk vi et grunnlag for å identifisere hvilke bedrifter i bransjen som benytter seg av de respektive evalueringsformene, og dermed egnet seg som studieobjekter. I tillegg leste vi aktuelle nyhetsoppslag og artikler.

Basert på bakgrunnsintervjuer og nyhetsoppslag laget vi en oversikt over aktuelle bedrifter, og sendte åtte av dem en henvendelse via e-post. Der presenterte vi oss selv, temaet og formålet med studien. Det var vesentlig å understreke at både bedriftene og intervjuobjektene ville bli behandlet anonymt, da dette ellers kunne forhindret deltagelse.

Etter henvendelsene ble sendt ut var det lite positiv respons i starten. Derfor fulgte vi opp henvendelsene med telefonsamtaler og opplevde da at flere bedrifter ønsket å delta i studien. Bedriftene syntes temaet virket spennende, samtidig som de anså studien nyttig og relevant å delta i. På bakgrunn av studiens tidsbegrensning, valgte vi å fokusere på tre bedrifter og tilhørende respondenter. Etter telefonsamtalene avtalte vi intervjudatoer og sendte respondentene informasjon om hensikten med utredningen. I henhold til Saunders et al. (2012) kan slik informasjon påvirke studiens validitet og reliabilitet, da respondentene kan sette seg inn i temaet i forkant av intervjuene. I tillegg kan informasjonen gi oss som intervjuere kredibilitet (Saunders et al., 2012).

Enheter

Vi ønsket å sammenligne bruk av skjønnsmessige prestasjonsevalueringer og tvungen fordeling i ulike bedrifter i oljebransjen. De tre studieobjektene ble valgt på bakgrunn av hvorvidt de var egnet til å besvare problemstillingen. Det var derfor sentralt at bedriftene brukte både skjønn og tvungen fordeling i prestasjonsevalueringer, og evalueringen måtte være tilknyttet ansattes avlønning.

Alle bedriftene vi har intervjuet er operatørselskaper i oljebransjen. Bedrift A er et internasjonalt olje- og gasselskap med hovedkontor i Norge. Både bedrift B og C er norske enheter i større internasjonale olje- og gasselskaper. I Norge har bedrift A flere ansatte enn de to andre bedriftene. Videre har bedrift C omtrent dobbelt så mange ansatte som bedrift B. Av anonymitetshensyn velger vi å ikke gå nærmere inn på informasjon om bedriftene. I stedet vil vi presentere systemene for prestasjonsevalueringer og belønning inngående i kapittel 4.

Intervjuobjekter

Valg av intervjuobjekter kan ha stor betydning for analysen, og det er dermed viktig å foreta en gjennomtenkt utvelgelsesprosess (Johannessen et al., 2011). Når vi baserer datainnsamlingen på dybdeintervjuer er vi avhengig av å basere oss på et utvalg, da intervjuer med samtlige ansatte vil være urealistisk. I følge Saunders et al. (2012) er det ikke nødvendigvis slik at man får mer utfyllende resultater fra hele populasjonen, sammenlignet med sampling fra et utvalg som representerer denne populasjonen. Enkelte argumenterer for at sampling muliggjør høyere sannsynlighet for presisjon, da man har mulighet til å benytte mer tid på hvert enkelt objekt (Henry, 1990).

Problemstillingen er sentrert rundt å finne ut hvilke muligheter og utfordringer som oppstår når prestasjonsevalueringer baseres på skjønn og tvungen karakterfordeling. Vi er derfor avhengig av å intervjue både ledere med evalueringsansvar og ikke-ledende ansatte. Ledere gir oss mulighet til å oppnå forståelse for utformingen og formålet med systemene, samt muligheter og utfordringer de møter ved prestasjonsevalueringer. Ikke-ledende ansatte vil gi oss innsikt i ansattes holdninger til å bli evaluert og motta karakterer som grunnlag for prestasjonsbasert lønn.

I alle bedriftene ønsket vi å snakke med minimum to ledere og to ansatte. For oss var det viktig at en av lederne jobbet innenfor HR, da de kunne bidra med særlig verdifull informasjon vedrørende belønningssystemer og evalueringsprosesser. I mangel på intervju med HR-leder i bedrift A gjennomførte vi et intervju med en av bedriftens eksperter på evalueringsprosesser, som sitter ved et kontor i utlandet. Utenom ønsket om en HR-leder i hver bedrift var målet å skaffe heterogene intervjuobjekter som kunne bidra med ulike refleksjoner og perspektiver. Basert på dette fremla bedriftene forslag til respondenter som kunne bidra med relevant informasjon til studien.

Vår empiriske undersøkelse er basert på 15 dybdeintervjuer der hvert intervju varte mellom en og to timer. Dette er innenfor Saunders et al. (2012) sin anbefaling om 5-25 enheter når man benytter semistrukturert form. Alle respondentene har til felles at de er høyt utdannet og har tilbrakt store deler av sin karriere i olje- og gassbransjen. De var svært kompetente og egnet seg godt til å svare utdypende på våre spørsmål. Vi intervjuet totalt ti menn og fem kvinner. Vi opplever at intervjuobjektene skilte seg fra hverandre, noe som ga oss et bredere perspektiv på studiens tema.

Intervjuobjektene hadde et spenn i alder fra 29 til 64 år. I bedrift A intervjuet vi en ekspert på evalueringsprosesser, en av lederne innenfor strategi og forretningsutvikling, en leder innen finans, samt tre ikke-ledende ansatte med ingeniørbakgrunn. I bedrift B intervjuet vi HR-leder, IT-leder, en leder innen finans og to ikke-ledende ansatte. Av ikke-ledende ansatte intervjuet vi en økonom og en ingeniør. I intervjuet med IT-leder stilte vi spørsmål fra både intervjuguiden rettet mot ledere, og fra intervjuguiden rettet mot ikke-ledende ansatte, fordi vi hadde en overvekt av lederintervjuer i bedriften. I bedrift C intervjuet vi HR-leder, teknisk leder, og to ikke-ledende ansatte. Av ikke-ledende ansatte intervjuet vi en økonom og en ingeniør. Følgende tabell gir en oversikt over intervjuobjektene i de tre bedriftene.

	Bedrift A	Bedrift B	Bedrift C
Ledere med evalueringsansvar	Ekspert på evalueringsprosesser, kvinne	HR-leder, kvinne	HR-leder, kvinne
	Leder innenfor strategi og forretningsutvikling, mann	IT-leder, mann	Teknisk leder, mann
	Finansleder, mann	Finansleder, kvinne	
Ikke-ledende ansatte uten evalueringsansvar	Ingeniør, mann	Økonom, mann	Økonom, kvinne
	Ingeniør, mann	Ingeniør, mann	Ingeniør, mann
	Ingeniør, mann		

Tabell 2: Intervjuobjekter

3.2.3 Intervjuguide

Et semistrukturert intervju gir mulighet for løsere dialog mellom intervjuer og respondent, men tar utgangspunkt i en intervjuguide (Saunders et al., 2012). Ghauri og Grønhaug (2010) hevder intervjuguiden skal ta utgangspunkt i studiens problemstilling og relevant forskningslitteratur, og fungere som en mal for intervjuene. Samtidig som vi benyttet eksisterende teori om skjønnsmessige prestasjonsevalueringer og tvungen fordeling, var vi åpne for at disse teoriene ikke var uttømmende. Spørsmålene måtte således både kunne avdekke i hvilken grad respondentenes opplevelser stemte med teori, samt gi mulighet til å avdekke ny innsikt i studiens tema. Ved å utforme en intervjuguide fikk vi sikret at spørsmålene var dekkene til problemstillingen (Saunders et al., 2012). I tillegg gjorde godt forarbeid det enklere å vite når oppfølgingsspørsmål var nødvendig, og når vi trengte mer detaljer. Vi fikk dermed mulighet til å skaffe ny innsikt ved å bygge videre på spørsmålene i intervjuguiden, avhengig av hva respondentene svarte.

Vi utformet to ulike intervjuguider, en til ledere med evalueringsansvar og en til ikke-ledende ansatte. Begge intervjuguidene ble delt inn etter relevant teori, men med tanke om at samtalen skulle flyte mest mulig naturlig. Fullstendig intervjuguide finnes i vedlegg 9.1. Ved gjennomføring av intervjuene ønsket vi innledningsvis å benytte åpne spørsmål om evalueringsprosessene og belønningssystemene for at respondentene kunne svare på egen måte, noe som gjør det mulig å avdekke ny innsikt på området (Saunders et al., 2012). Deretter stilte vi spørsmål som utforsket særlig relevante elementer for vår problemstilling ved deres svar.

Vi oppdaget tidlig i intervjuprosessen at avvik fra intervjuguiden var nødvendig, da respondentene besvarte flere spørsmål i ett. Semistrukturerte intervjuer og en fleksibel og løs dialog var derfor egnet, da vi fikk tilpasset spørsmålene etter hvert som det var nødvendig. Dette var viktig for å få innhentet dybdeinformasjon og innsikt i nye aspekter gjennom oppfølgingsspørsmål. Intervjuguiden var likevel svært nyttig da den sikret fremdrift og systematikk i intervjuene, samt at vi fikk stilt de relevante spørsmålene vi hadde forberedt. En ulempe med å benytte semistrukturerte intervjuer er at intervjuene kan føres ulikt ettersom hvilke oppfølgingsspørsmål som blir naturlige og interessante å stille den enkelte vil variere. Dette kan gjøre sammenligningen i etterkant vanskeligere (Saunders et al., 2012).

3.2.4 Intervjuprosessen

I henhold til Saunders et al. (2012) er gode forberedelser en forutsetning for vellykkede intervjuer. Kunnskap om temaet og bedriftene man forsker på kan bidra til å skape troverdighet i intervjusituasjonen og medføre at intervjuobjektene gir mer detaljerte svar. Før intervjuene brukte vi derfor mye tid på å lese oss opp på bransjen, bedriftene og temaet vi forsker på. For innsikt i bransjen leste vi avisoppslag og artikler tilgjengelig for allmennheten, samt gjennomførte bakgrunnsintervjuer med tre personer med erfaring fra bransjen og bedriftene vi skulle intervjue.

Samtlige intervjuer ble gjennomført ved bedriftenes lokaler i henholdsvis Stavanger og Oslo. Dette gjorde vi av tre ulike grunner. For det første ønsket vi å bruke minst mulig av intervjuobjektenes tid, både for å vise respekt for deres tid og fordi en annen løsning kunne forhindret deltakelse. Videre har Saunders et al. (2012) vist at lokasjon er en viktig faktor for hvorvidt intervjuobjektene føler seg komfortable. For det tredje vil intervjuer som avholdes ansikt til ansikt gi rikere informasjon, da man her får mulighet til å observere kroppsspråk og reaksjoner (Saunders et al., 2012).

I henhold til Saunders et al. (2012) vil begynnelsen av intervjuet være avgjørende for en vellykket datainnsamlingen. Dette skyldes at det er de første minuttene man bygger troverdighet og dermed kan legge et godt grunnlag for resten av intervjuet. Vi innledet hvert intervju med å presentere oss selv og temaet for oppgaven. Her forklarte vi også enkelte viktige begreper for å sikre at respondentene forstod spørsmålene på en best mulig måte. Videre presiserte vi både bedriften og intervjuobjektene sin anonymitet for å legge til rette for en åpen og ærlig samtale. Intervjuobjektene ble også fortalt at det ble respektert om de valgte å avstå fra enkelte spørsmål, slik at de ikke skulle føle seg presset på noen måte. Ingen av intervjuobjektene benyttet seg av denne muligheten.

Det var ideelt for oss at intervjuene kunne tas opp på bånd, da man ved bruk av båndopptaker har en komplett form for registrering av rådata ved kvalitative undersøkelser (Jacobsen, 2015). Videre gir bruk av båndopptaker intervjuer mulighet til å fokusere på å stille gode oppfølgingsspørsmål og følge med i samtalen, fremfor å ta notater (Saunders et al., 2012). Gode oppfølgingsspørsmål er avgjørende når man gjennomfører en eksplorerende studie. Samtlige intervjuobjekter godkjente bruk av båndopptaker. Vi ga dem mulighet til å

få tilsendt transkriberingen i etterkant for å sikre at det ikke var noen misforståelser, men ingen ønsket dette.

Selve intervjuet innledet vi med å stille spørsmål om intervjuobjektene sin stilling for å få nødvendig bakgrunnsinformasjon, men også for å sikre en myk start. Deretter stilte vi åpne spørsmål relatert til problemstillingen. Åpne spørsmål bidro til at respondentene i størst mulig grad kunne svare etter egne erfaringer og resonnement, uten å bli ledet inn på spor (Saunders et al., 2012). Slike spørsmål er også hensiktsmessig da det gir respondenten mulighet til å svare mer utdypende, noe som kan avdekke holdninger og fakta. Videre er det sentralt med tydelige og forståelige spørsmål for å redusere eventuelle bias og øke reliabiliteten til studien (Saunders et al., 2012). Vi opplevde det derfor viktig å være i stand til å reformulere spørsmålene og sikre at respondentene forstod oss. Oppfølgingsspørsmål var derfor med på å heve kvaliteten på forskningen og unngå misforståelser (Saunders et al., 2012).

Ved alle intervjuene var begge tilstede. Dette ønsket vi da studiens validitet styrkes om man i etterkant av intervjuene kan diskutere hva respondentene mente med sine utsagn. Vi ville begge skulle observere kroppsspråk, ordlegging, stemningen ved spørsmålene og andre forhold som er av betydning for hvordan et utsagn burde tolkes, men som ikke fanges opp av båndopptakeren. Det var viktig at begge observerte, da den ene kan ha gjort observasjoner den andre gikk glipp av. I henhold til Kvale og Brinkmann (2009) er forholdene rundt intervjuet svært viktig, da et utsagn i seg selv kan tolkes på ulike måter.

Samtlige intervjuer varte mellom en og to timer. Vi avklarte med hvert enkelt intervjuobjekt hvor lang tid de skulle sette av, og var nøye med å holde oss innenfor avtalt tidsramme. Vi gjennomførte intervjuene 28.september, 4.oktober og 31.oktober i bedrift A, den 12.oktober i bedrift B og 20.oktober i bedrift C.

3.3 Analyse av data

3.3.1 Klassifisering av data

Studiens eksplorerende forskningsdesign og innhenting av kvalitative primærdata legger grunnlag for analysen. For å forenkle analyseprosessen startet vi med å transkribere samtlige båndopptak, altså gjorde den muntlige talen skriftlig. Dette bidro til at vi fikk oversikt over

og ble kjent med datamaterialet (Saunders et al., 2012). I tillegg til hva som ble sagt skrev vi ned stemningen ved spørsmålene, ordlegging, kroppsspråk og andre faktorer av betydning for hvordan utsagnene burde tolkes. Videre fokuserte vi på spørsmålene vi stilte for å se om det var eventuelle ledende spørsmål, spørsmål som ga grunnlag for misforståelser og om spørsmålene ble stilt på samme måte til de ulike intervjuobjektene. Transkribering var tidkrevende, men verdifullt både for oss og andre. Det var verdifullt for oss ved at vi fikk reflektert over intervjuobjektenes utsagn og fikk ideer til hvilke faktorer som var av interesse for videre analyse. Videre gjør transkriberingen det mulig for andre å undersøke våre tolkninger i ettertid (Jacobsen, 2015).

Etter transkriberingen analyserte vi data gjennom en templatanalyse. Saunders et al. (2012) hevder dette er hensiktsmessig ved kvalitative data, særlig når forskningstilnærmingen er abduktiv. En templatanalyse er nyttig for å velge ut viktige temaer som burde utforskes videre (Saunders et al., 2012). Prosessen i en templatanalyse består av å utvikle kategorier og koble disse til utdrag fra dataene (Saunders et al., 2012). Dette er en fleksibel metode som tilrettelegger for forhåndsdefinering av koder. Kodene kan senere modifiseres og man kan legge til nye koder etter hvert som data samles og analyseres, og nye sammenhenger og temaer avdekkes (Saunders et al., 2012). Kodene vi benyttet gjorde det enklere å sortere datamaterialet i kategorier, og gjorde dataene mer forståelige og håndterlige. Videre ga kategoriene grunnlag for å se sammenhenger og utvikle påstander. Vi testet påstandene mot alternative forklaringer, fordi godt testede påstander gir grunnlag for mer valide og velbegrunnede konklusjoner (Saunders et al., 2012).

3.3.2 Presentasjon av funn

I kapittel 4 presenteres funn vedrørende utformingen av bedriftenes systemer for prestasjonsevaluering og belønning. Deretter presenteres relevante funn fra templatanalysen i kapittel 5, der sentrale funn illustreres gjennom sitater. Basert på empiriske funn vil vi i kapittel 6 sammenligne og diskutere på tvers av de tre bedriftene opp mot teori og empiri, for å gi en utvidet forståelse for problemstillingen.

3.4 Evaluering av metode

I det følgende vil vi evaluere kvaliteten på studien. For dette formål er det hensiktsmessig å vurdere egenskapene ved det eksplorerende forskningsdesignet gjennom fokus på validitet og reliabilitet.

3.4.1 Validitet

"*Validitet i kvalitative undersøkelser dreier seg om i hvilken grad forskerens funn på en riktig måte reflekterer formålet med studien og representerer virkelighet*" (Johannessen et al., 2006, s.199). Generelt blir casestudier ofte kritisert for lav grad av **begrepsvaliditet**, da subjektive evalueringer brukes for å samle inn data. Yin (2014) definerer begrepsvaliditet som hvorvidt måleenhetene som benyttes reflekterer aspektene forskerne har til hensikt å studere. For å minimere sjansen for feiltolkninger og fremme meningsfulle funn, forklarte vi begreper og uttrykk tydelig, og forsikret oss om at intervjuobjektene forstod meningen av disse før vi stilte spørsmål (Saunders et al., 2012). Videre benyttet vi samme intervjuguide til alle ledere med evalueringsansvar, og samme intervjuguide til ikke-ledende ansatte. Dette gjorde vi for å sikre likebehandling, god struktur og et mer sammenlignbart datamateriale.

Intern validitet dreier seg om hvorvidt funnene oppfattes som riktige, om beskrivelsene er sanne og om sammenhenger er reelle (Jacobsen, 2015). Her gjelder også om forskerne har fått tilgang til respondentenes kunnskap og erfaringer, samt tolker utsagn slik intervjuobjektene faktisk har ment dem (Saunders et al., 2012).

For å styrke den interne validiteten har vi gjort en rekke tiltak i intervjuprosessen. Vi har intervjuet personer på flere nivåer i samtlige bedrifter, og tilegnet oss god kunnskap om tvungen fordeling og skjønnsmessige prestasjonsevalueringer gjennom intervjuer med ledere med evalueringsansvar. Ved intervjuer med ikke-ledende ansatte har vi fått konkrete og personlige erfaringer vedrørende opplevelser rundt karakter basert på skjønn og tvungen fordeling. For å være i best mulig stand til å forstå intervjuobjektenes faktiske erfaringer og kunnskap benyttet vi åpne spørsmål og oppfølgingsspørsmål aktivt. Vi mener derfor vi har skaffet innsikt i både kontekst og fenomen, noe som styrker den interne validiteten (Saunders et al., 2012).

Til tross for intervjuer med ønsket antall og type intervjuobjekter, vil det fortsatt være et spørsmål om intervjuobjektene har gitt sann informasjon. Når det gjelder kildene til informasjon burde man ha større tiltro til de som er så nær fenomenet som mulig (Jacobsen, 2015). På bakgrunn av dette har vi valgt å basere studien på data direkte fra ansatte i bedriftene og hvordan de opplever skjønnsmessige prestasjonsevalueringer og tvungen fordeling. For et bredere perspektiv har vi intervjuet både ledere med evalueringsansvar og ikke-ledende ansatte. Vi har derfor skreddersydd intervjuguidene til henholdsvis ledere og ikke-ledende ansatte, som beskrevet i avsnitt 3.2.3, for å sikre at intervjuobjektene fikk spørsmål som var nærliggende å svare på i deres stillinger. For å undersøke forskningens troverdighet er en mulighet å sammenligne studiens funn med annen forskning (Saunders et al., 2012). Det er lite forskning på nøyaktig samme problemstilling som inneværende studie, men teori og forskning på lignende problemstillinger viser tilsvarende funn. Dette styrker den interne validiteten.

I intervjuprosessen har det vært viktig for oss at spørsmålene som ble stilt var i naturlig rekkefølge inndelt etter temaer, og tydelige for intervjuobjektene. Vi forklarte derfor viktige begreper etter hvert som vi byttet tema. For å kunne stille gode og gjennomtenkte spørsmål satte vi oss grundig inn i bedriftene i forkant av intervjuene, noe som bidro til styrket kredibilitet overfor intervjuobjektene (Saunders et al., 2012). Grundig forarbeid var fordelaktig da intervjuobjektene var i stand til å gi mer utdypende svar, samtidig som vi var bedre egnet til å tolke svarene i etterkant.

Jacobsen (2015) hevder at ”*informasjon fra flere uavhengige kilder gir en gyldig beskrivelse av fenomenet* ” (s. 231). Vi har på bakgrunn av dette valgt å intervjue samtlige intervjuobjekter individuelt for å sikre at kildene var uavhengige av hverandre. Dette bidrar til å styrke den interne validiteten.

Den interne validiteten er også avhengig av om vi har tolket dataene på en hensiktsmessig måte og om det gis en troverdig representasjon av data. Vi har som beskrevet i avsnitt 3.3.1. hatt en grundig prosess med behandling av data, samtidig som vi har en kritisk fremstilling i analysen. Vi har i tillegg sammenlignet fellestrekk ved intervjuobjektenes utsagn, noe som reduserer sannsynligheten for å trekke feilaktige slutninger. Slike tiltak styrker den interne validiteten.

Vi tilbød samtlige intervjuobjekter å få tilsendt transkriberingen, for å kunne avdekke eventuelle feiltolkninger. Tilbakeføring og bekreftelse av resultatene kan styrke validiteten. Det var imidlertid ingen som ønsket å benytte seg av muligheten, noe som medfører at vi kan ha feiltolket data. På den andre siden unngikk vi med dette at intervjuobjektene modererte eller trakk tilbake utsagn.

Ekstern validitet, eller generaliserbarhet, omhandler hvorvidt funnene i studien kan gjøres gjeldende for lignende sammenhenger (Saunders et al., 2012). Kvalitativ forskning baseres på få enheter og til et spesifikt formål, noe som trekker i retning av lav overførbarhet. Det vil således være begrensninger for grad av generalisering (Jacobsen, 2015), særlig med tanke på statistisk generalisering. En kan likevel få teorietisk eller analytisk generalisering som kan sannsynliggjøres, men ikke bevises. Det vil i henhold til Williams (2004) derfor kun være moderat grad av overførbarhet i en slik studie. Imidlertid er ikke hensikten med denne eksplorerende studien å generalisere funn, men få dypere innsikt i prestasjonsevalueringer basert på skjønn og tvungen fordeling. Funnene fra forskningen kan ikke generaliseres, men heller gi indikasjon på hvilke utfordringer og muligheter som oppstår når skjønnsmessige prestasjonsevalueringer kombineres med tvungen fordeling.

Til tross for enkelte svakheter konkluderer vi med tilfredsstillende grad av validitet i studien.

3.4.2 Reliabilitet

Reliabilitet viser til hvorvidt tilsvarende funn ville blitt avdekket om studien ble gjentatt eller utført av en annen forsker (Saunders et al., 2012). Reliabiliteten sier således noe om påliteligheten ved forskningsdesignet, og om resultatene er konsistente på bakgrunn av datagrunnlaget. Generelt blir casestudier ofte kritisert for lav grad av reliabilitet, da funnene vil være et resultat av realiteten på forskningstidspunktet og derfor utsatt for endringer (Saunders et al., 2012). Imidlertid er ikke nødvendigvis formålet med casestudier å kunne replikeres, men heller reflektere virkeligheten på tidspunktet data ble samlet inn (Saunders et al., 2012).

I casestudier vil reliabiliteten i stor grad avhenge av datainnsamlingen (Yin, 2009). **Intervjuerbias** er når intervjueren handler på en måte som påvirker intervjuobjektet sine svar. Denne skjevheten skyldes ofte uklare svar (Saunders et al., 2012), noe vi prøvde å unngå ved å stille oppfølgingsspørsmål i tilfeller vi var usikre på respondentenes forklaringer. I slike situasjoner gjenfortalte vi svaret de avga slik vi forstod det, og fikk

bekreftet eller avkreftet om innholdet var forstått slik det var ment. Intervjuerbias kan også forekomme dersom intervjuer har subjektiv oppfattelse av svarene, eller stiller ledende spørsmål. Vi fokuserte derfor på å stille åpne spørsmål uten føringer. I etterkant av intervjuene har vi diskutert uklare utsagn og sammenhenger for å sikre best mulig forståelse.

En trussel mot reliabiliteten kan være at ledere svarer på spørsmålene for å fremme bedriften, og ikke-ledende ansatte svarer det de tror lederne ønsker de skal svare. Selv om både intervjuobjekter og bedrifter behandles anonymt i studien, kan intervjuobjektene frykte deres meninger kan spores tilbake til dem. Våre funn og meningsytringer tyder imidlertid på at dette ikke har vært tilfelle i utstrakt grad.

Grad av reliabilitet vil også påvirkes av **responsbias**. Responsbias foreligger når intervjuobjektet påvirkes av omgivelser eller formulering av spørsmålene (Saunders et al., 2012), og kan medføre feilaktige svar. For å unngå denne biasen sørget vi for at intervjuobjektene følte seg trygge og avslappede, hadde intervjuene ved bedriftenes kontorer og satte av god tid. Vi innledet intervjuene med en rolig samtale for å sikre trygghet i intervjusituasjonen. Videre presiserte vi anonymiteten til hver enkelt, for å fremme ærlige og åpne svar. Dette var særlig viktig da personlige opplevelser og erfaringer er sensitivt for intervjuobjektene, samtidig som systemer for evaluering og belønning av ansatte er sensitivt for bedriftene ovenfor media og konkurrenter. Imidlertid er det ikke mulig å garantere at responsbias unngås, da vi ikke vil være i stand til å vite om de fortalte sannheten om sine opplevelser og meninger.

Videre kan **deltakerbias** påvirke studiens reliabilitet. Deltakerbias er når man ikke kan kontrollere hensikten til de som aksepterer å delta i studien (Saunders et al., 2012). En mulighet er at bedriftene eller intervjuobjektene som sier seg villig til å delta har en agenda med å stille opp. I så måte kan studien ha et skjevt utvalg. Imidlertid viser våre funn en nyanserikdom som man vanskelig ville avdekket ved utstrakt deltakerbias.

Båndopptak av intervjuene kan være en mulig kilde til redusert reliabilitet i vår studie. Ved opptak av intervjuene risikerer man mer tilbakeholdne intervjuobjekter. Vi var derfor nøye med å ikke legge opptaksutstyret synlig foran intervjuobjektene, for å unngå å stadig minne intervjuobjektene på båndopptaket. Vi opplevde intervjuobjektene som stadig mer komfortable utover i intervjuet, og vi stilte derfor de mest personlige og sensitive spørsmålene mot slutten. Videre vil transkriberingen av intervjuene styrke studiens

reliabilitet, da vi unngikk informasjonstap og subjektive tolkninger (Saunders et al., 2012). Transkriberingen i kombinasjon med notater om stemning, kroppsspråk og andre forhold ved intervjuene mener vi styrker studiens reliabilitet.

Det er med eksplorerende forskningsdesign urealistisk å anta at alle trusler mot reliabiliteten er unngått. Likevel mener vi å ha møtt utfordringene på en grundig og hensiktsmessig måte, og konkluderer derfor med tilfredsstillende grad av reliabilitet i studien.

3.5 Etiske hensyn

Ghauri & Grønhaug (2010) definerer etikk som de moralske prinsipper og verdier som påvirker hvordan forskeren utfører de ulike delene av sin studie. I forskningsarbeidet er det vesentlig å handle ut fra et etisk riktig perspektiv, da mye tillit vies forskeren (Saunders et al., 2012). Vårt forskningsdesign har innebåret etiske utfordringer vi har tatt hensyn til og forsøkt å minimere. Forskningen må ikke medføre bedriftene som deltar noen ulemper, og de må oppleve å behandles med respekt og ærlighet. I forkant av intervjuene ga vi derfor tilstrekkelig informasjon til bedriftene og intervjuobjekter slik at de kunne ta en vel overveid beslutning om de ønsket å delta eller ikke.

Saunders et al. (2012) presiserer hvor viktig det er å ivareta rettighetene til de som deltar i studien. Ivaretakelse av bedriftene og intervjuobjektene sin anonymitet har derfor vært sentralt. Som bakgrunn for deres deltakelse var anonymiteten presisert allerede ved første kontakt, noe som var særlig viktig med tanke på studiens sensitive tema. Anonymiteten kan ha bidratt til å skape tillit, og derfor mer åpen dialog og informasjonsdeling. Videre var det viktig å presisere at intervjuobjektene kunne velge å trekke seg fra intervjuet, eller avstå fra enkelte spørsmål dersom det var ønskelig. Deltakerne kunne ellers følt begrensninger når de besvarte enkelte spørsmål, noe som kunne påvirket responsen som er grunnlaget for våre funn og konklusjoner (Saunders et al., 2012). Ingen av intervjuobjektene ønsket å avstå fra spørsmål eller trekke seg, noe som tyder på trygghet i intervjusituasjonen.

I etterkant av intervjuene måtte vi bruke innsamlet informasjon på en hensynsfull måte, og funn ble rapportert korrekt, nøyaktig og i sin helhet (Saunders et al., 2012). Opptakene fra intervjuene ble slettet umiddelbart da vi ikke lenger hadde bruk for dem, noe vi også avtalte med intervjuobjektene for at de skulle føle seg sikret. Ved analyse og bearbeidelse av data

sikret vi deltakernes anonymitet, ved eksempelvis å unngå sitater som lett kunne kobles til enkeltpersoner.

3.6 Oppsummering av studiens metode

I dette kapittelet har vi beskrevet og argumentert for valg av metodisk tilnærming til studien. Valg av metode er gjort med hensikt å besvare vår problemstilling om prestasjonsevalueringer basert på skjønn og tvungen karakterfordeling. Studien har et eksplorativt formål, og en abduktiv tilnærming og kvalitativ metode er derfor egnet. For å sikre nyanserik dybdeinformasjon gjennomførte vi semistrukturerte intervjuer i tre ulike bedrifter. Tabellen under oppsummerer studiens metode.

Studiens metodiske tilnærming	
Forskningsdesign	Eksplorativt formål
Forskningstilnærming	Abduktiv tilnærming
Forskningsmetode	Kvalitativ metode
Forskningsstrategi	Multippel casestudie
Datainnsamling	Semistrukturerte dybdeintervjuer
Antall case	3
Antall intervjuer	15 i bedriftene, 3 bakgrunnsintervjuer
Lengde på intervjuer	1-2 timer
Analyse av data	Transkribering og templatanalyse
Validitet	Tilfredsstilt
Reliabilitet	Tilfredsstilt

Tabell 3: Oppsummering av metodisk tilnærming

4. Bransjen og casene

I dette kapittelet gjør vi kort rede for bransjen og bedriftene vi har undersøkt. Deretter presenteres funn vedrørende bedriftenes systemer for belønning og prestasjonsevaluering.

4.1 Bransjen og casene

Denne studien tar for seg tre bedrifter i den norske oljebransjen. Gjennom våre bakgrunnsintervjuer har vi blitt fortalt at bransjen er preget av kompetansesterke ansatte, der det stadig er fokus på å videreutvikle og forbedre deres prestasjoner. Flere av bedriftene har derfor innført omfattende systemer for prestasjonsevalueringer. Evalueringene summeres ofte i karakterer som forventes å passe inn i forhåndsbestemte fordelinger. Ansattes karakterer knyttes typisk til ulike former for prestasjonsbasert avlønning, blant annet kollektive bonuser, individuelle bonuser, prestasjonsbaserte lønnsjusteringer og prestasjonsbaserte aksjer.

Oljebransjen har tidligere vært preget av høy lønnsomhet og vekst, men har siden oljeprissjokket i 2014 opplevd nedgang i lønnsomhet (Martén & Whittaker, 2015). Nedgangstidene har resultert i nedbemanninger (Aadland, 2015) og nulloppgjør i lønnsøkninger og bonuser i flere oljerelaterte selskaper.

De tre bedriftene vi har undersøkt er operatørbedrifter i bransjen. Bedriftene vil heretter bli omtalt ved navnene “bedrift A”, “bedrift B” og bedrift “C”. Bedrift A er et internasjonalt olje- og gasselskap med hovedkontor i Norge. Både bedrift B og C er norske enheter i større internasjonale olje- og gasselskaper. Rangert etter antall ansatte i Norge er bedrift A størst. Videre er bedrift C større enn bedrift B. Av anonymitetshensyn unngår vi å presentere bedriftene nærmere, men fokuserer heller på deres systemer for prestasjonsevaluering og belønning. Alle bedriftene benytter karakterer i prestasjonsevalueringene. Karakterene er knyttet til ulike former for prestasjonsbasert avlønning.

4.2 Bedriftenes systemer for belønning og prestasjonsevaluering

I det følgende presenteres bedriftenes systemer for belønning og prestasjonsevalueringer. Vi fokuserer på delene av belønningssystemene som er tilknyttet karakterene til ikke-ledende ansatte. Ledere med evalueringsansvar vil gjennomgående bli omtalt som ledere, og ikke-ledende ansatte uten evalueringsansvar vil bli omtalt som ansatte. Først presenteres systemene til bedrift A, deretter bedrift B og avslutningsvis bedrift C. Til slutt oppsummeres funnene i en tabell.

4.2.1 Bedrift A

Bedrift A sitt belønningssystem

Overordnet ønsker bedrift A å tilby konkurransedyktig lønn, men ikke være lønnsledende. Dette skal bidra til å tiltrekke og beholde talenter i bedriften. Bedriften benytter prestasjonsbaserte komponenter for å motivere til innsats, med bakgrunn i at økt prestasjon skal gi økt økonomisk uttelling.

Bedrift A benytter fastlønnsordning med prestasjonsbaserte lønnsøkninger. Størrelsen på den prestasjonsbaserte lønnsøkningen avhenger av karakteren den enkelte får i den årlige prestasjonsevalueringen. Forskjellen i lønnsøkning mellom de sterkeste og svakeste yterne utgjør vanligvis et par prosentpoeng, men kan i særlig gode tider være opp til 10 prosentpoeng. Allokeringen av lønnsøkning må imidlertid være innenfor rammene gitt av forhandlinger med fagforeningene. Det innebærer at hver gruppe kollektivt skal få gjennomsnittlig lønnsøkning på en fastsatt prosent. Som følge av nedgangstidene har de prestasjonsbaserte lønnsøkningene blitt fjernet de siste to årene, lønnen har kun blitt justert for inflasjon.

I tillegg til lønnsøkninger får alle ansatte konsernbonus. Størrelsen på konsernbonusen avgjøres av bedriftens resultater og resultatene til forretningsområdet man tilhører. I praksis har det imidlertid blitt gitt samme konsernbonus til alle ansatte uavhengig av forretningsområde siden systemet ble introdusert. Ansatte på ledernivå og oppover har også tilgang til individuell bonus. Det individuelle bonussystemet er gradert, i form av at prosentvis bonus blir større jo høyere man kommer opp i systemet. Bonusen blir også

korrigert i henhold til konjunkturer i bransjen. Eksempelvis har bonusmulighetene vært redusert som følge av nedgangstidene.

Bedrift A sitt system for prestasjonsevalueringer

Lønnsøkninger og individuelle bonuser blir i bedrift A bestemt på bakgrunn av skjønnsmessige prestasjonsevalueringer lederne foretar av sine medarbeidere. De skjønnsmessige vurderingene kommer til uttrykk gjennom karakterer fra 1 til 5, der 5 er beste karakter. Bedriften bruker en form for det vi har definert som tvungen fordeling av ansattes karakterer. Det forventes både at hver leders gjennomsnittlige karaktergivning skal ligge på 3, og normalfordeling av ansattes karakterer. Imidlertid blir ikke denne fordelingen overholdt i praksis, da 95 % av ansatte får en kombinasjon av karakterene 3 og 4.

Den enkelte medarbeider får to karakterer, en på leveranse og en på atferd. Leveransemålene er typisk mer kvantifiserbare, og nært knyttet til den enkelte medarbeiders stillingsbeskrivelse. Eksempelvis kan leveransemål innebære tidsfrister, budsjettrammer eller antall prosjekter som skal fullføres i løpet av året. Atferdsmål er nærmere knyttet til bedriftens verdier, eksempelvis samarbeidsevne, initiativtaking og kunnskapsdeling. Slike mål skal bidra til godt arbeidsmiljø og velfungerende team. De to karakterene på henholdsvis leveranse og atferd vektlegges likt ved fastsettelse av finansielle belønninger. En leder forteller om lik vektlegging av atferd og leveranse slik:

"Vi legger lik vekt på de to karakterene når de ansattes belønning bestemmes. Om du for eksempel har en score på 5-4 eller 4-5 har ingenting å si i forhold til kroner og øre. De er helt likestilt i praksis."

Gjennomsnittskarakteren 3 gis til ansatte som leverer som forventet. Får man karakterer over 3 leverer man mer enn det som forventes, mens karakterer under 3 indikerer at man har levert under forventet og må forbedre sine prestasjoner. Videre har bedriften en glidende karakterskala, der forventningene til den enkelte justeres over tid. For hvert lønnsnivå er det bygget opp forventninger til hva og hvordan man skal levere. Å levere som forventet og få karakteren 3 blir derfor vanskeligere for ansatte på høyere lønnsnivå, og jo lenger man har jobbet i bedriften. Dette underbygges av en leder vi har snakket med:

"For hvert lønnsnivå har vi på en måte bygget opp en forventning. Når du kommer ung inn så har du mye å lære, derfor lar vi oss lettere imponere hvis du gjør en god jobb, og det er en litt lavere terskel for å levere bra. (...) Etterhvert som folk stiger i gradene så øker også

forventningene. Har man allerede fått økt lønnen sin gjennom å hoppe lønnstrinn opp, så betyr dette at man må legge mer igjen. Det kreves da vesentlig skarpere og bedre leveranser enn hos de på lavere lønnsnivå."

I den årlige medarbeidersamtalen setter ansatte sammen med sin nærmeste leder mål på leveranse og atferd for kommende år. Det er imidlertid en dynamisk prosess, der målene kan endres underveis ved behov. I løpet av året skal ledere evaluere ansattes prestasjoner og måloppnåelse. Ledere skal ikke kun evaluere basert på egne observasjoner, men gjennom året hente tilbakemeldinger fra andre ledere og interessenter med innblikk i ansattes prestasjoner. Basert på egne og andres observasjoner setter leder veiledende karakterer til sine ansatte. Karakterene hver leder setter i sin gruppe, rapporteres videre opp til de øverste lederne innenfor hvert forretningsområde. Disse lederne vil da gjennomføre en kalibreringsprosess. Kalibreringen innebærer at lederne går gjennom og justerer de veiledende karakterene for å sikre normalfordeling og gjennomsnittskarakter på 3. Det er her de endelige karakterene som gis i medarbeidersamtalene blir satt. Disse karakterene gis det ikke mulighet til å klage på. En leder utdyper kalibreringsprosessen i følgende sitat:

"Karakteren 5 sitter veldig høyt hos oss. Kanskje det bare er en eller to i teamet som skal få karakteren 5, slik at man få finne ut hvem man føler har gjort det best i løpet av året i forhold til forventningene. Det innebærer at en ansatt kanskje vil få karakteren 4 fordi vi ikke kan gi flere 5´ere. De måles altså relativ opp mot hverandre. (...) Det kan for eksempel være en person jeg gir karakteren 5, men som jeg vet at til nøds kan forsvare en 4'er om det blir nødvendig i kalibreringen. Da skriver jeg dette som et notat til de som gjennomfører kalibreringen for mitt forretningsområde. Det er sånn sett ikke sikkert at karakterene jeg gir som veiledende blir stående etter kalibreringsprosessen, fordi ansatte på tvers av ulike grupper måles mot hverandre."

4.2.2 Bedrift B

Bedrift B sitt belønningssystem

HR-leder i bedrift B forteller at de har et veldig raust belønningssystem og tilbyr høyere lønninger og bonuser enn de fleste konkurrenter i Norge. De ønsker med dette å rekruttere og beholde medarbeidere med høy kompetanse, samt motivere ansatte til å yte maksimalt. En leder forklarer det på følgende måte:

"Det belønningssystemet vi har tror jeg ikke du finner maken til, det har vi gjort benchmark på. Det er et veldig raust belønningssystem. Korttidsperspektivet er å være konkurransedyktig på lønn så vi tiltrekker de beste, og det langsiktige målet er å beholde de beste. (…)Formålet med belønningssystemet er å få de ansatte til å yte mer og gjøre en ekstra god jobb. Vi ønsker jo at bonussystemet skal gi de ansatte incentiv til å yte maksimalt, og belønne de som virkelig gjør det"

Bedrift B opererer med et fastlønnssystem. Alle ansatte har mulighet for en individuell prestasjonsbasert lønnsøkning som avhenger av karakteren den enkelte får i den årlige prestasjonsevalueringen. Som følge av nedgangstidene har de hatt nulloppgjør i lønnsøkninger de siste årene.

I tillegg til prestasjonsbaserte lønnsøkninger er samtlige medarbeidere omfattet av et individuelt bonussystem, der bonusen er en prosent av fastlønnen. Hvor stor prosentvis bonus man får avhenger av karakter i prestasjonsevalueringen og bedriftens overordnede resultater. Karakteren avgjør øvre og nedre grense for hvor stor bonusen kan være, mens bedriftens overordnede resultater avgjør den prosentvise bonusen man mottar innenfor dette intervallet. Det er store forskjeller mellom bonuser tilknyttet de ulike karakterene, der forskjellen mellom to karakterer kan være flere hundre tusen. Nedgangstidene har ikke påvirket bonusordningen i utstrakt grad, da ansatte har fått 85 %- 95 % av bonusen de ville fått i normalår med tilsvarende karakterer.

For nøkkelpersoner legges det i tillegg inn langsiktige insentiver i belønningssystemet. Da vil en gitt prosentandel av bonusen bli gjort om til aksjer, der halvparten kan utbetales etter to år, og den resterende summen etter tre år. Målet er at de langsiktige insentivene skal medvirke til å beholde særlig viktig kompetanse i bedriften.

Bedrift B sitt system for prestasjonsevalueringer

Bedrift B baserer lønnsøkninger og bonuser på skjønnsmessige prestasjonsevalueringer lederne foretar av sine ansatte. De skjønnsmessige vurderingene kommer til uttrykk gjennom et karaktersystem som består av fire kategorier. Kategoriene er verbale og delt inn etter hvorvidt man presterer "under forventet", "som forventet", "over forventet" eller om prestasjonen er "fremragende". Bedriften bruker en form for tvungen fordeling av ansattes karakterer, med fastsatte prosentandeler for hvor mange som kan få hver karakter.

Hovedvekten av de ansatte, omtrent 70-80 %, skal få karakteren "som forventet". Kun et fåtall skal få hver av de andre karakterene.

Den enkelte ansatt får karakter både på leveranse og atferd, der de to karakterene samles til en enkelt karakter når belønningen skal fastsettes. Leveranse er knyttet til hvordan den ansatte har prestert rent faglig, mens atferd knytter seg til personlig egnethet. Da bedriften synes det interessante er kombinasjonen av hva som leveres og hvordan det gjøres, vektlegges leveranse og atferd likt i evalueringsprosessen. En leder forklarer hensikten med å måle både leveranse og atferd slik:

"Det er viktig for oss å evaluere ansatte både på leveranse og atferd. Forskjellen når man evaluerer også på atferd er at man får vurdert om ansatte samarbeider, er flink til å dele informasjon, om de bidrar til det sosiale og om de er gode til å jobbe i team. Ved at man evalueres på begge områder sikrer man at ansatte jobber på lag og at man ikke får medarbeidere som konkurrerer. Det er flere aspekter enn leveranse som er viktig for å være en god medarbeider, og ved å måle på begge områder sikrer man at de andre forholdene også blir prioritert."

Det er laget klare retningslinjer for hva som skal kreves for å få de ulike karakterene på både leveranse og atferd. I tillegg finnes det retningslinjer for hvordan disse to karakterene skal samles til en helhetlig karakter. Om en ansatt får karakteren "som forventet" på enten leveranse eller atferd, er hovedregelen at den ansatte skal få "som forventet" i samlet karakter. Dette gjelder både om den andre karakteren man får er "under forventet" og "over forventet". For å få "over forventet" i samlet karakter må man derfor ha denne karakteren på både atferd og leveranse. Tabellen under illustrerer hvordan de to karakterene skal samles til en karakter.

LEVERANSE ATFERD	Under forventet	Som forventet	Over forventet	Fremragende
Fremragende	Som forventet	Over forventet	Fremragende	Fremragende
Over forventet	Som forventet	Som forventet	Over forventet	Fremragende
Som forventet	Som forventet	Som forventet	Som forventet	Over forventet
Under forventet	Under forventet	Som forventet	Som forventet	Som forventet

Tabell 4: Karakterer på leveranse og atferd

I bedriften er det ikke et system for hvordan ledere skal justere krav til ansatte over tid. Hvorvidt krav justeres over tid avhenger av hvilken leder man har.

Ansatte og ledere fastsetter sammen ved begynnelsen av hvert år mål for det kommende året. Her går de også gjennom hva som kreves for å få de ulike karakterene. I tillegg har bedriften en midtveisevaluering halvveis i året, der målene justeres hvis det er behov for det. De har midtveisevaluering fordi målene som burde prioriteres kan endre seg i løpet av året, og det anses derfor hensiktsmessig med en dynamisk prosess. Ved årsslutt evaluerer nærmeste leder, med eventuelle innspill fra andre om nødvendig, hvordan mål på både leveranse og atferd er nådd. Lederen gir deretter sine ansatte en veiledende karakter i medarbeidersamtalen. Etter medarbeidersamtalene rapporteres de veiledende karakterene videre til de øverste lederne i bedriften som gjennomfører en kalibreringsprosess. De ansatte måles da relativt mot hverandre og endelige karakterer blir satt. Kalibreringen sikrer at karakterene er innenfor den forhåndsbestemte fordelingen. Ansatte gis ikke mulighet til å klage på de kalibrerte karakterene. HR-leder forklarer begrunnelsen for kalibreringsprosessen på følgende måte:

"De ansatte måles relativt mot hverandre fordi det er en fordelingskurve som det forventes at ansattes karakterer skal passe inn i. Det innebærer at en ansatt kanskje vil få karakteren "som forventet" fordi vi ikke har flere "over forventet" å gi ut."

4.2.3 Bedrift C

Bedrift C sitt belønningssystem

Overordnet ønsker bedrift C å tilby konkurransedyktig lønn som skal bidra til å tiltrekke og beholde talenter i bedriften. Videre ønsker de at prestasjonsbaserte elementer ved belønningssystemet skal motivere til innsats og stimulere til videreutvikling av stadig bedre prestasjoner. Bedriften ser på prestasjonsbaserte belønninger som viktig i kunnskapsbedrifter som gjør det bedre om enkeltpersoner og grupper presterer godt. Samtidig som de ønsker å belønne ekstraordinære prestasjoner, ønsker de å sikre en følelse av fellesskap og unngå for store lønnsdifferanser. Oppsummert forklarer HR-leder formålet med belønningssystemet slik:

"Hovedformålet med vårt belønningssystem er å belønne de som yter mest, og at de skal fortsette med samme engasjement videre. Selvsagt er det også viktig at vi klarer å tiltrekke og beholde talenter (...). Vi ønsker å ha et differensiert system slik at vi kan belønne ekstraordinær innsats, men samtidig sikre at forskjellene ikke blir for store. En balanse mellom ytelsesbasert lønn og en viss likhet i lønninger er viktig for oss."

Bedrift C opererer med et fastlønnssystem med individuelle prestasjonsbaserte lønnsøkninger hvert år. Lønnsøkningen avhenger av karakteren ansatte får i den årlige prestasjonsevalueringen, men den må være innenfor en årlig budsjettramme og forhandlinger med fagforeningene. Kollektivt skal hver gruppe få en gjennomsnittlig lønnsøkning på en fastsatt prosent forhandlet frem med fagforeningene, men lederne kan gi mer eller mindre til ansatte med karakterer over eller under gjennomsnittet. I tillegg til lønnsøkninger får alle ansatte konsernbonus. Konsernbonusen avgjøres på bakgrunn av ulike mål, som eksempelvis ulykkesstatistikk, hvorvidt bedriften holder seg innenfor budsjett og bedriftens resultater.

Bedrift C opererer videre med et system som kalles "eksepsjonelle belønninger". Dette systemet innebærer spontane bonuser ved fremragende prestasjoner eller hvis ansatte har nådd en milepæl. Slike bonuser gis sjeldent og har ingenting med den årlige lønnen å gjøre, men er en bonus du får der og da ved særlig god innsats. I tillegg har bedriften prestasjonsbaserte aksjer som deles ut samtidig som eventuelle lønnsjusteringer.

Bedrift C sitt system for prestasjonsevalueringer

I bedrift C blir lønnsøkninger og utdeling av aksjer bestemt av skjønnsmessige prestasjonsevalueringer ledere foretar av sine medarbeidere. De skjønnsmessige

vurderingene kommer til uttrykk gjennom karakterer fra 1 til 5, der 1 er beste karakter. Bedriften bruker en form for det vi har definert som tvungen fordeling av ansattes karakterer. Det er en forventning fra hovedkontoret at karakterene skal passe inn i en normalfordelt kurve. Samtidig foreligger en forventning om gjennomsnittskarakter på 3. Forventning om gjennomsnitt og normalfordeling gjelder både for hver enkelt leder sin gruppe og bedriften som helhet.

Størrelsen på lønnsøkningen knyttes til karakteren man får. Dersom en ansatt får karakteren 3, har man levert som forventet og gjort en tilfredsstillende jobb. Da vil man få lønnstillegg som forhandlet frem med fagforeningene. Leverer man dårligere enn dette vil det derfor gis et lavere lønnstillegg enn gjennomsnittet, men om man leverer bedre gis relativt mer. Bedriften har ikke et system for hvordan ledere skal justere krav til ansatte over tid. Hvorvidt krav justeres over tid avhenger av hvilken leder man har.

Den enkelte ansatt får karakter både på leveranse og atferd, der de to karakterene samles til en enkelt karakter når belønningen skal fastsettes. Leveransemål er typisk mer kvantifiserbare og nært knyttet til stillingsbeskrivelsen den enkelte medarbeider har. Atferdsmål er nærmere knyttet til bedriftens verdier, hvordan ansatte fungerer i stillingen og i samhandling med andre.

Ansatte og ledere fastsetter sammen ved begynnelsen av hvert år mål som skal gjelde det kommende året. I tillegg har de en midtveisevaluering halvveis i året, der målene justeres hvis det er behov for det. Midtveisevalueringen anses hensiktsmessig fordi hvilke mål som burde prioriteres kan endre seg i løpet av året. Ved årsslutt evaluerer nærmeste leder og dens overordnede, med eventuelle innspill fra andre om nødvendig, hvordan mål på både leveranse og atferd er nådd. Nærmeste leder og dens overordnede setter så en veiledende karakter som sendes videre til kalibrering. Kalibreringsprosessen foretas av HR-avdelingen og noen av bedriftens øverste ledere før karakterene gis i medarbeidersamtalene. Kalibreringen innebærer at lederne går gjennom og justerer de veiledende karakterene for å sikre gjennomsnittskarakter på 3 og ønsket fordeling.

Bedriften har en sikkerhetsventil i systemet for prestasjonsevalueringer, der ansatte har mulighet til å klage på karakteren mottatt av nærmeste leder. Dersom en ansatt ikke føler seg riktig oppfattet eller observert, kan de be om vurdering fra andre de mener har bedre innsikt i

deres prestasjoner. HR-avdelingen er da pliktig å skaffe tilbakemeldinger fra de det gjelder, og ta klagen til etterretning. Klagestatistikken i bedriften er lav.

4.2.4 Oppsummering av bedriftenes systemer for belønning og prestasjonsevaluering

I dette kapittelet har vi presentert bedriftenes systemer for belønning og prestasjonsevalueringer. Alle bedriftene har prestasjonsbaserte belønninger tilknyttet karakterene ansatte får i medarbeidersamtalen. Karakterene forventes å tilfredsstille forhåndsbestemte fordelinger, men hvilke fordelinger som benyttes, og hvor strengt ledere må forholde seg til fordelingene, varierer mellom bedriftene. Tabellen under oppsummerer bedriftenes systemer for belønning og prestasjonsevalueringer.

	Bedrift A	Bedrift B	Bedrift C
Prestasjonsbaserte lønnsøkninger	Ja, direkte tilknyttet ansattes karakterer	Ja, direkte tilknyttet ansattes karakterer	Ja, direkte tilknyttet ansattes karakterer
Prestasjonsbasert bonus knyttet til karakterene	Nei, ikke for ansatte som ikke er på ledernivå	Ja, for samtlige ansatte	Nei, men har prestasjonsbaserte aksjer og eksepsjonelle belønninger
Karakterer	Ja, skala fra 1-5	Ja, fire verbale karakterer	Ja, skala fra 1-5
Tvungen fordeling	Ja, i form av forventning om gjennomsnittskarakter og normalfordeling	Ja, fastsatte prosentandeler som kan få hver karakter. Kravet er ufravikelig	Ja, i form av forventning om gjennomsnittskarakter og normalfordeling
Kalibrering	Ja, før medarbeidersamtalen	Ja, etter medarbeidersamtalen	Ja, før medarbeidersamtalen
Karakter på atferd og leveranse	Ja	Ja	Ja
Justering av krav over tid	Ja	Ikke system for justering av krav	Ikke system for justering av krav
Antall som evaluerer	Minimum 1. Nærmeste leder. Skal også hente innspill fra andre	1.Nærmeste leder. Kan hente innspill fra andre	2. Nærmeste leder og dens overordnede. Kan hente innspill fra andre
Klagemulighet	Nei	Nei	Ja

Tabell 5: Oppsummering av bedriftenes systemer for belønning og prestasjonsevaluering

5. Funn til studiens problemstilling

I dette kapittelet presenteres relevante funn for vår problemstilling:

Hvordan virker prestasjonsevalueringer basert på skjønn og tvungen fordeling av ansattes karakterer?

Først presenteres funn fra bedrift A, deretter bedrift B og til slutt bedrift C. For hver bedrift oppsummeres sentrale funn i en tabell. Vi vil gjennomgående i dette kapittelet omtale ledere med evalueringsansvar som ledere, og ansatte uten evalueringsansvar som ansatte.

5.1 Funn i bedrift A

I det følgende presenteres funn fra bedrift A som er relevante for å besvare studiens problemstilling. Grunnet gjennomført ekspertintervju i bedrift A har vi mer informasjon om denne bedriften enn de to andre bedriftene. For å sikre en systematisk gjennomgang vil vi først presentere funn som omhandler skjønnsmessige prestasjonsevalueringer, før vi presenterer funn vedrørende karakterer og tvungen fordeling.

5.1.1 Skjønnsmessige prestasjonsevalueringer i bedrift A

Bedrift A bruker skjønn i prestasjonsevalueringene av ansatte, noe som innebærer både muligheter og utfordringer. En gjennomgående trend er at tillit til leders prestasjonsevaluering er avgjørende for opplevelsen av skjønnsmessige vurderinger. Ansattes tillit til leders prestasjonsevaluering avhenger av hvorvidt leder tar hensyn til alle relevante aspekter ved deres arbeidshverdag, har mulighet til å observere innsatsen de legger ned, har erfaring med evalueringsprosesser, og forholdet mellom dem.

Helhetlige vurderinger

I bedrift A opplyser samtlige ledere at muligheten for å gi helhetlige vurderinger av de ansatte er hovedårsaken til at de bruker skjønnsmessige prestasjonsevalueringer. Bruk av skjønn er nødvendig om ledere skal kunne gi en god og tilpasset vurdering av de ansatte, fordi ansatte i bedriften ofte har komplekse og sammensatte arbeidsoppgaver. Følgelig er det krevende å utvikle mål som er grundige nok til å fange opp alle relevante aspekter ved den enkeltes arbeidshverdag. Skjønnsmessige vurderinger av ansattes prestasjoner er dermed

hensiktsmessig for å få med nyanserikdommen i de ansattes arbeidshverdag. En leder forklarer formålet med å benytte skjønn på følgende måte:

"Skjønnsmessige vurderinger er nødvendige fordi vi ikke klarer å lage målene grundige nok, eller det ville vært veldig tidkrevende å gjøre dem perfekte. (...). Skjønnet må være der for at man skal kunne se den enkelte, så unik som hver og en av oss er, og få folk til å yte maksimalt. Det er kun med skjønn du kommer helt innpå den enkelte."

Som nevnt i del 4.2.1 får ansatte karakter både på leveranse og atferd, der begge karakterene har lik betydning for deres belønning. I henhold til ledere er skjønnsmessige vurderinger av atferd egnet for å sikre evaluering av viktige ikke-kvantifiserbare elementer ved ansattes arbeidshverdag. Slik ønsker bedriften å hindre nedprioritering av viktige aspekter ved arbeidshverdagen som er vanskelig å måle objektivt. En annen viktig grunn til å evaluere ansatte på atferd er å sikre at de handler i tråd med bedriftens verdier, og å signalisere at positive bidrag til arbeidsmiljøet verdsettes. For bedriften er det særlig viktig å tilrettelegge for godt samarbeidsklima, da samarbeid i mange tilfeller er avgjørende for ansattes prestasjoner. Derfor måles ansatte på hvorvidt de bidrar til godt samarbeid og gjør andre gode. Ansatte ser på sin side på samarbeid som en selvfølge i deres arbeidshverdag. Hensikten med å evaluere samarbeidsevner illustreres av en leders utsagn:

"Skjønnet vi utøver går blant annet på om vi er tilfreds med måten de ansatte har opptrådd sammen med andre. Jeg har gått igjennom dette med å faktisk samarbeide og dele kunnskap, det skal alle ha på sin agenda. Hvis ikke risikerer jeg for mange individualister som jager sin egen vinning (...) Poenget er at alle skal være med på å gjøre hverandre gode, og det premierer vi for."

Ukontrollerbare forhold

Prestasjonene til de ansatte i bedriften blir ofte påvirket av eksterne omgivelser. Ledere og ansatte ser derfor på skjønnsmessige vurderinger som positivt da det gir mulighet til å ta høyde for ukontrollerbare forhold. Samtlige ansatte forteller at de opplever evalueringsprosessen dynamisk, og at lederne tar hensyn til støy som påvirker deres prestasjoner. Målet med å ta hensyn til ukontrollerbare forhold er å øke ansattes motivasjon også i nedgangstider hvor objektive måltall ikke nødvendigvis reflekterer den enkeltes innsats. En leder utdyper hvordan det tas hensyn til ukontrollerbare forhold i dette sitatet:

"I karakterene tar vi hensyn til det hvis det har oppstått problemer eller hendelser utenfor den enkelte ansattes kontroll. Det er viktig at vi viser at vi ser innsatsen de legger ned, selv når ikke resultatene er der. Hvis ikke vil jo motivasjonen i disse tider gå kraftig nedover. Vi prøver derfor å bli flinkere til å vise gjennom karakterene at vi ser det når folk legger inn ekstra innsats. (...) Men vi må også se litt på hvordan ansatte responderer på slike hendelser. Ofte har de delvis mulighet til å dempe negative effekter eller utnytte positive effekter av støy, og da er det viktig at det evalueres."

På en annen side forteller ledere og ansatte at støy som påvirker ansattes prestasjoner positivt ikke filtreres ut på samme måte som negativt støy. Ansatte opplyser blant annet om tilfeller der de belønnes for vellykkede prosjekter selv når det skyldes flaks heller enn ekstraordinær innsats. Enkelte opplever dette negativt, da det for dem virker urettferdig når andre som yter tilsvarende innsats får høyere belønning på grunn av flaks. En ansatt forklarer dette slik:

"Jeg ser at noen som har vært heldige å få bli med på et svært vellykket prosjekt får veldig gode karakterer, selv om dette prosjektet kanskje bare varte i to uker. Andre som har jobbet like hardt eller hardere gjennom hele året kan ende med en helt grei karakter. Sånn sett kan det av og til oppleves litt urettferdig."

Observasjon av ansattes innsats

Samtlige ansatte forteller at kvaliteten på prestasjonsevalueringene avhenger av lederens mulighet til å observere den enkeltes innsats. I mange tilfeller jobber ansatte og ledere tett sammen. Nærmeste leder har da god innsikt i de ansattes arbeidshverdag. I andre tilfeller er det større avstand mellom ledere og ansatte. Ledere er da pålagt å hente innspill fra andre med bedre innsikt i ansattes prestasjoner. Imidlertid forteller ledere og ansatte om enkelte utfordringer knyttet til å basere prestasjonsevalueringer på andres observasjoner. Slike tilbakemeldinger innebærer ofte mindre nyanser, og er derfor dårligere egnet til å gi et helhetlig bilde av ansattes innsats. I tillegg krever det mye tid og ressurser av lederen å hente inn informasjonen. En ansatt forteller om lederes evne til å observere ansattes innsats på følgende måte:

"Jeg synes det er vesentlig at lederen som evaluerer meg faktisk er i stand til å sette seg inn i min arbeidshverdag og har god innsikt i hva jeg holder på med. Det er jo min nærmeste leder som kjenner meg best og evaluerer meg (...). Jeg har opplevd flere ganger at min leder har snakket med andre for å få mer informasjon om hvordan jeg har jobbet, og det synes jeg

er bra. Det som er ulempen i sånne situasjoner er at tilbakemeldingene leder får fra andre gjerne er mindre innholdsrike."

Når ledere henter informasjon fra flere som har observert ansatte i løpet av året, gis et bredere bilde av ansattes prestasjoner. Da ledere mener at prestasjonsevalueringer basert på informasjon fra flere er mer uangripelige, gis det ikke mulighet på å klage på slike karakterer.

Tilbakemeldinger

Forståelse og innsikt i ansattes prestasjoner er også viktig for leders evne til å gi gode tilbakemeldinger gjennom året. En ansatt forteller at hvorvidt tilbakemeldingene de mottar er nyttige og motiverende avhenger av om leder setter seg inn i deres arbeidsoppgaver. Uten nødvendig kunnskap om ansattes arbeid, blir tilbakemeldingene fort for generelle og dermed lite prestasjonsfremmende.

Eksperten på prestasjonsevalueringer forteller at bedriften ikke har sterk nok kultur for tilbakemeldinger. Eksperten mener karakterene og evalueringene ofte får for stort fokus, noe som fortrenger de jevnlige og gode tilbakemeldingene. Enkelte ansatte uttrykker et ønske om flere meningsfylte tilbakemeldinger gjennom året, men de fleste opplever sine ledere som flinke til å gi positive og konstruktive tilbakemeldinger. Tilbakemeldingene oppmuntrer dem til å gjøre en enda bedre jobb. En ansatt beskriver dette slik:

"Generelt synes jeg ledere er flinke til å gi tilbakemeldinger, og jeg synes jeg får tilbakemeldinger som både tydeliggjør hvordan jeg kan forbedre meg, og ros når det er på sin plass. Noen ganger synes jeg likevel at de kunne hatt enda mer fokus på jevnlige tilbakemeldinger gjennom året."

Formålet med jevnlige tilbakemeldinger er i følge en leder å sikre at ingenting virker overraskende i den årlige medarbeidersamtalen. Dersom en ansatt har gjort en mindre god jobb er ledere opptatt av å kommunisere dette klart, for å gi ansatte mulighet til å forbedre prestasjoner og unngå dårlig karakter ved årsslutt. Imidlertid forteller de her om forskjeller mellom ledere, der enkelte kvier seg for å gi kritiske tilbakemeldinger. Samtalene kan bli vanskelige å ta fordi noen ansatte er mindre mottakelige for kritikk. Enkelte ledere unngår derfor kritiske tilbakemeldinger de frykter kan gå på bekostning av relasjonen de har til sine ansatte. En leder forteller om utfordringen med kritiske tilbakemeldinger som følgende:

"Hvis de blir veldig overrasket over karakteren de får, har jeg gjort noe feil i forhold til forventningsavklaringen. For eksempel har jeg innført her at ingen skal bli overrasket når de får en dårlig karakter, de skal få en advarsel og mulighet til å rette opp i løpet av året(...) Men jeg ser jo eksempler på ledere som ikke tør å gi de dårlige tilbakemeldingene. Det er ikke alle ansatte som takler kritikk så godt fordi de har et selvbilde som ikke samstemmer helt med virkeligheten. Men det er veldig viktig at de tør å ta de samtalene og gi de kritiske tilbakemeldingene, slik at lederen har gjort det den kan for å forbedre sine ansattes prestasjoner."

Trynefaktor

Ved skjønnsmessige prestasjonsevalueringer vil det alltid være en risiko for at forhold som ikke burde være relevante for karaktersettingen blir vektlagt i vurderingen, bevisst eller ubevisst. Blant annet kan evalueringene påvirkes av relasjonen mellom leder og ansatt. Samtlige ledere og ansatte forteller om forholdet mellom leder og ansatt som avgjørende for hvorvidt prestasjonsevalueringene oppleves rettferdige. Statistikken i bedriften viser generelt positive dialoger mellom ledere og ansatte, men eksperten på prestasjonsevalueringer forteller også om situasjoner med konflikter og misnøye ved karaktersetting.

Misnøye knyttes blant annet til trynefaktor i prestasjonsevalueringene. Både ledere og ansatte forteller om risiko for at trynefaktor kan påvirke vurderingene, da det alltid er enkelte man har bedre kjemi med enn andre. Hvorvidt trynefaktor oppleves som et problem avhenger i følge ansatte av lederen man har, da enkelte ledere lettere lar seg påvirke av hvem de liker bedre ved karaktersetting. Verken ledere eller ansatte opplever imidlertid trynefaktor som et utstrakt problem.

En utfordring når trynefaktor kan påvirke prestasjonsevalueringer, er i følge ledere at enkelte ansatte prøver å smiske for å få bedre karakter. De viser til situasjoner der ansatte har prøvd å påvirke dem, men oppfatter smisking enkelt å gjennomskue, sjeldent og uproblematisk.

Både karakteren på atferd og leveranse settes basert på skjønnsmessige vurderinger, og vektlegges likt i prestasjonsevalueringen. Samtlige ledere i bedriften forteller imidlertid at det er mer utfordrende å evaluere og følge opp ansattes atferd enn leveranser, noe som knyttes til mindre tydelige atferdsmål. Ansatte opplever det også vanskeligere å bli vurdert på atferd enn leveranser, da uklare mål medfører mer uhåndgripelige evalueringer. Som følge av krevende atferdsevalueringer, forteller en leder at vurderingen av ansattes leveranser

ofte legger føringer for vurderingen av atferd. Om en ansatt eksempelvis har levert dårlig på leveranse skal det derfor mye til å få en høy score på atferd. Lederen forklarer det slik:

"Det er mye vanskeligere å evaluere ansattes atferd enn leveranser. Dette er både fordi det er mindre klare retningslinjer for oss som evaluerer atferden, men også fordi vi vet at ansatte synes det er krevende å få tilbakemelding på atferd. (...). Atferdsiden blir nok litt formet av leveransesiden. Har en ansatt levert veldig godt på leveranse, skal det derfor mye til å få en veldig dårlig score på atferd."

Motvilje mot å differensiere

Ledere forteller at det for dem er viktig å bevare gode relasjoner til sine ansatte. De synes det er ubehagelig å gi ansatte svake karakterer, da de frykter det kan påvirke relasjonen til ansatte negativt. Dette gjelder selv når en svak karakter er velbegrunnet, da ansattes persepsjon av egne prestasjoner ofte er høyere enn det som er reelt. De nevner blant annet eksempler på situasjoner der ledere har gitt bedre karakterer enn fortjent for å unngå misnøye og konflikter. En leder utdyper det slik:

"Mange ledere opplever det vanskelig å gi ansatte svake karakterer fordi de ønsker å unngå konflikter og ubehagelige samtaler. Jeg har sett eksempler hvor ledere skyver unna problemet ved å gi ansatte karakterene 3-3 i stedet for å ta den konfrontasjonen. Dette er en negativ effekt, da det man egentlig gjør er å skjule problemet. I tillegg gjør det at det ikke skilles tilstrekkelig mellom gode og mindre gode prestasjonene, så de det går utover er de som faktisk gjør en god og tilfredsstillende jobb. (...) Å ta de samtalene kan være ubehagelig på kort sikt, men på lenger sikt er det fornuftig og nødvendig."

5.1.2 Karakterer og tvungen fordeling i bedrift A

Bedrift A bruker tvungen fordeling av ansattes karakterer. De har både en forventning om normalfordeling og et krav om gjennomsnittskarakter på 3, noe som innebærer både muligheter og utfordringer. I det følgende vil vi presentere funn vedrørende karakterer og tvungen fordeling.

Faglig og personlig utvikling

Et av formålene med karaktersystemet er å motivere til faglig og personlig utvikling. Det skal fungere som et oppfølgingsverktøy, gjennom å gi veiledning til de ansatte basert på mål

som samsvarer med bedriftens verdier og forretningsutfordringer. Dette illustreres av en leders sitat:

"Det vi ønsker med karaktersystemet er å spore folk an til å utvikle seg, og presse seg selv til å gjøre mer enn hva de typisk er komfortable med.(...) Også har man sagt at det at du utvikler deg og det at du faktisk responderer positivt på det bedriften ber deg om, det skal du få igjen for."

Imidlertid forteller eksperten på evalueringsprosesser i bedrift A om for sterkt fokus på karakterer og selve vurderingen, når det sentrale burde være å skaffe innsikt i hvordan den enkelte medarbeider kan forbedre ytelse. Et ønske er derfor å fjerne bruken av karakterer, og heller konsentrere seg om å gi innsikt i hvordan den enkelte kan forbedre seg gjennom jevnlige og gode tilbakemeldinger.

Dokumentasjon og tilbakemeldinger

Flere ledere nevner at bruk av karakterer innebærer omfattende dokumentasjon av prestasjonsevalueringene som ligger til grunn for belønningene. Ansatte mener dokumentasjonen gjør det mer synlig hvorfor de får den belønningen de gjør. På den måten oppleves prestasjonsevalueringene mer objektive og rettferdige, og legger til rette for en mer korrekt lønnsutvikling. Eksperten på evalueringsprosesser i bedrift A sier et av formålene med å gi ansatte karakterer nettopp er å skape en transparent link mellom belønning og ytelse, og gjøre evalueringene med håndgripelige for de ansatte.

Et annet formål med bruk av karakterer i prestasjonsevalueringene er å skape åpenhet gjennom tilbakemeldinger fra ledere og andre kollegaer. Systemet stimulerer til dialog mellom ansatte og ledere, og gir mulighet for å gi verdifulle og pålitelige tilbakemeldinger på innsatsen. Videre sikres minimum en formell samtale i året. En leder forklarer det på følgende måte:

"Et annet formål med å gi karakterer er at det gir ledere et rammeverk som er godt forankret rundt hvordan de kan gi tilbakemeldinger. Systemet tvinger ledere til å søke tilbakemeldinger fra andre ledere og interessenter, og sikrer minst en samtale per år."

Imidlertid mener en annen leder at kalibrering av karakterene gjør det vanskeligere å gi gode og meningsfulle tilbakemeldinger i løpet av året. Ledere kan være nervøse for å gi

prestasjonstilbakemeldinger som ikke er kalibrert, da tilbakemeldingene kan misforstås som en karakterindikasjon av de ansatte.

Ytelsesstandarder og identifisering

Eksperten på evalueringsprosesser sier et av formålene med kalibrering på tvers av grupper og avdelinger er å muliggjøre en delt forståelse for hva hver karakter betyr med tanke på ytelsesstandarder. Kalibreringen tvinger ledere på tvers av avdelinger til å kommunisere rundt karaktersetting, og slik sikre at de evaluerer etter samme kriterier. For å sikre like ytelsesstandarder har bedriften i tillegg blitt mer opptatt av å sette klare retningslinjer for hvordan karakterene skal settes, og i større grad styre skjønnet. Imidlertid er det i følge eksperten på prestasjonsevalueringer likevel forskjeller i hvordan ledere gir karakterer og utøver skjønn. En leder forklarer det slik:

"I bedriften er det veldig mange ledere, og da er det nok ikke å forvente at alle tenker likt. Da oppstår det fort at noen avdelinger er veldig rause, mens andre er ekstremt konservative. Det er klart at det er kjedelig for ansatte som jobber for en mindre raus leder og ser at andre grupper lettere oppnår gode karakterer."

Et annet formål med kalibrering av ansattes karakterer er å identifisere sterke, normale og svake ytere. Identifiseringen kan bidra i administrative beslutninger, som eksempelvis hvem som skal forfremmes og hvem som har utviklingsbehov. Karakterer kan på en annen side virke demotiverende, noe en leder forklarer på følgende måte:

"Å ha en merkelapp med hvilken karakter man har kan senke årlig ytelse, og mennesker er motvillige mot å få en merkelapp på den måten. Det kan være ganske demotiverende. (...) Vi har hatt ansatte som har blitt sykemeldt og til og med sluttet grunnet misnøye rundt karaktersetting."

Differensiering

Alle ledere er fra sentralt hold gitt beskjed om å bruke hele karakterskalaen, og måle de ansatte relativt mot hverandre. Ledere som faktisk bruker skalaen forteller om tilfeller de blir fristet til å gi karakterer etter hva som passer i kalibreringen, heller enn karakterer som samsvarer med faktiske prestasjoner. Et eksempel kan være når et team som helhet har ytt signifikant over eller under gjennomsnittet. I slike tilfeller burde egentlig alle i teamet fått tilsvarende karakterer om de skal reflektere faktiske prestasjoner. Når systemet likevel

tvinger ledere til å måle medarbeiderne relativt mot hverandre støtter det ikke opp under teamytelse.

Til tross for at ledere er gitt beskjed om å bruke skalaen, er det få ledere som faktisk gjør det. En av hovedutfordringene med prestasjonsevalueringene i bedrift A er nettopp liten grad av differensiering mellom gode og dårlige prestasjoner. Selv om differensiering kan virke negativt på teamytelse, kan mangel på differensiering innebære andre negative effekter. Samtlige ledere og ansatte ytrer derfor et ønske om større grad av differensiering. Bedriften har et veldig teoretisk system for å skille og rangere ansattes prestasjoner, men det blir ikke brukt tilstrekkelig i praksis. Eksperten på evalueringsprosesser viser til karakterfordelingen der hele 95 % av alle ansatte i bedriften får en kombinasjon av karakterene 3 og 4. Eksperten tviler på at kun 5 % av ansatte i bedriften yter mindre eller signifikant mer enn det som forventes.

Ledere forteller at det svært krevende å bruke skalaen, og unngår derfor å normalfordele karakterene. Dette knyttes til at det er utfordrende i en kompetansesterk bedrift å identifisere hvem som skal få de laveste karakterene, da det er få medarbeidere med svake prestasjoner. Med forventning om gjennomsnittskarakter på 3 kan en leder ikke gi en ansatt karakter over forventning uten å samtidig gi en annen ansatt karakter under forventning. Når det da er vanskelig å identifisere svake ytere gis det heller ikke mulighet til å premiere gode prestasjoner. Samlet fører derfor forventningen om gjennomsnittskarakter på 3 til lav grad av differensiering. De ansatte forstår utfordringen, og en av dem beskriver det på følgende måte:

"Det er vanskelig å differensiere i vår gruppe da vi er mange individer med mye erfaring og høy kompetanse. Jeg skjønner ikke hvordan man da kan gi noen i vår gruppe karakteren 1 eller 2 (...). Jeg har forstått at systemet har en kalibrering, og om man da skal gi noen karakteren 5, må noen få en svakere karakter. Jeg skjønner det da blir vanskelig å dele ut de virkelig gode karakterene. (...) Dette er nok en utfordring for lederne, men det er likevel synd at systemet gjør det vanskelig å premiere gode prestasjoner."

Samtlige ansatte gir uttrykk for at differensiering i karakterer er avgjørende om karakterene skal virke motiverende. De opplever det urettferdig når svært mange med ulike prestasjoner ender med tilsvarende karakterer. I tillegg erfarer de at karakterene er like fra år til år, selv om prestasjonene endrer seg. En ansatt forklarer det slik:

"Vi har som gruppe opplevd at uansett hva du gjør så kommer du ganske likt ut. Man får nesten karakteren 3-3 uansett, og det er lite motiverende. Det er synd, man har jo systemet for å bruke hele spekteret."

Gjennomsnittlig karakter på 3

Karakteren 3 er ment å være en god karakter, og der gjennomsnittet skal ligge i prestasjonsevalueringene. Begrunnelsen er at ansatte i bedriften rekrutteres fordi de gjerne besitter viktig erfaring og kompetanse. Om man da leverer som forventet er det fullt tilfredsstillende. Imidlertid opplyser ledere om utfordringer knyttet til å skaffe forståelse for dette, da ansatte opplever karakteren 3 demotiverende. En leder forteller om utfordringen på følgende måte:

"Jeg synes alltid det er kjedelig å måtte gi en score 3-3 til ansatte som har gjort en god jobb. Jeg er veldig fornøyd med jobben de gjør og ønsker at de skal fortsette å gjøre den samme jobben. Det som er kjipt er når jeg kommuniserer 3-3 så blir de skuffet, og hadde kanskje sett for seg en høyere score. Det det ender med er at folk som har gjort en god jobb blir demotiverte fordi karakteren 3 ikke oppfattes motiverende."

Dette underbygges av en av de ansattes svar da vi spurte om opplevelsen rundt å få karakteren 3:

"Vi har blitt fortalt at 3-3 er en veldig god evaluering. Og 3-3 er jo greit, men de fleste havner jo på 3'eren, så er det egentlig så bra? Er ikke 3 bare gjennomsnittlig da? Så sånn sett er det en litt demotiverende karakter å få hvis man har jobbet mye og levert godt."

Justering av krav over tid

Som nevnt i del 4.2.1. benytter bedriften en glidende karakterskala. De ønsker med dette å fremme kontinuerlig utvikling og stadig forbedrede prestasjoner. Dette gjelder for alle ansatte, men særlig for seniorer i bedriften som har kommet opp på høyere lønnsnivå. I følge eksperten på evalueringsprosesser legger et slikt system opp til inkonsistent målsetting, da grunnlaget for evaluering varierer ut fra ambisjonsnivå. For ansatte som har jobbet i bedriften over tid er det derfor ikke et sterkt insentiv til å være ambisiøs i individuell målsetting, både når det gjelder type og antall arbeidsoppgaver. En annen leder forteller om seniorer som mer motvillige mot å ta på seg krevende arbeidsoppgaver på grunn av høye forventninger, og forklarer det slik:

"På grunn av systemet får jeg noen suboptimaliseringer. De som er seniorer ønsker å ikke bli satt på vanskelige oppdrag, for da har jeg høyere forventning til dem. Selv om de vanskelige oppdragene ofte er noe de er veldig gode på, vil det være bedre for dem å få enklere prosjekter de kan gjøre det veldig bra på. Da får jeg plutselig en motsatt effekt av det som var meningen, nemlig ønsket om at de skal strekke seg."

Som konsekvens av økte krav over tid opplever de ansatte å få mindre og mindre igjen for økt innsats. De finner det nesten umulig å få høyere karakter enn 3 når forventningene som stilles er så høye. Ledere bekrefter at det er krevende for seniorer å få høyere karakterer enn 3. Ved kalibrering av ansattes karakterer prioriterer de ofte å gi nyutdannede de beste karakterene, da de gjerne utfører samme typer arbeidsoppgaver som seniorene, men til lavere lønn. Seniorene forstår begrunnelsen for å justere krav over tid, men opplever karaktersystemet likevel mindre motiverende. En ansatt forklarer det på følgende måte:

"Når man blir senior, er det nok sånn at karakterskalaen ikke blir brukt, det er et smalere spenn. Jeg føler vi ikke får igjen for innsatsen vi legger ned. Ledelsen forventer at vi leverer godt, og jeg vet det er sånn systemet er, men karakterene blir likevel mindre motiverende når man har jobbet her en stund."

Finansiell belønning

I bedrift A er prestasjonsevalueringen kun knyttet til lønnsøkninger, med mindre man er på ledernivå og har tilgang til individuell bonus. Formålet med å ha prestasjonsbaserte lønnsøkninger er å rettferdiggjøre forskjeller i arbeidsinnsats. Generelt er variasjonen i lønnsøkninger som gis basert på karakterene relativt små, men det kan i gode tider skilles opp til ti prosentpoeng. Det vanlige er imidlertid betydelig mindre forskjeller i lønnsjusteringer mellom de sterkeste og svakeste yterne. Både ansatte og ledere forklarer at utfordringen knyttet til for lite differensiering forsterkes når det ikke er store nok forskjeller i finansiell belønning tilknyttet ytelse. Ikke bare er det for lite differensiering, men de få som faktisk skiller seg ut i positiv eller negativ retning får ikke finansiell belønning som tydeliggjør forskjellen. Karakterene oppleves av enkelte lite motiverende når det ikke gis tilstrekkelig finansiell uttelling for gode prestasjoner. Systemet virker spesielt demotiverende for de flinkeste, da de vil få omtrent lik belønning som en som jobber mindre. I følge ledere og ansatte er det i ethvert insentivsystem viktig at de som gjør mest faktisk sitter igjen med størst finansiell belønning. Dette illustreres av en leders sitat:

"Jeg tror nok systemet sånn det er nå er demotiverende for de som leverer bra, og det er uheldig. Står du på ekstra skal du få uttelling for det. Det handler om motivasjon og anerkjennelse, og at man føler seg verdsatt når man gjør noe ekstra. Hos oss skulle vi gjerne hatt mer av det (...) Jeg tror ikke en situasjon sånn den er nå, der alle får tilnærmet det samme, virker motiverende."

Ønsket hos både ledere og ansatte er derfor større differensiering i karakterer, med tilhørende forskjeller i finansiell belønning. Flere ansatte og ledere stiller spørsmålstegn ved den egentlige nytten av systemet når det verken differensieres i karakterer eller lønn. En ansatt uttrykker utfordringen med for små forskjeller i finansielle belønninger på følgende måte:

"Jeg ser jo han som kommer sent og går tidlig og ikke legger inn noe ekstra, vi får akkurat det samme. Det skjer jo ikke noe. Det er litt demotiverende synes jeg. (..) Jeg stiller meg jo spørsmålet: hvor mye ressurser skal man bruke på måling og karaktersetting når utslagene på lønn og reward er såpass begrenset?. (...). Når man generelt har lønnsøkninger på 2-3-4%, og hvor forskjellen opp til toppscore kanskje bare er på 0,5 %, er det da verdt det å yte den ekstra innsatsen?"

En ytterligere utfordring med for små forskjeller i finansielle belønninger, er i følge ledere at de mister ansatte med toppkarakterer til konkurrenter som tilbyr høyere lønn. Slik systemet er i dag gis ledere for lite rom til å belønne ansatte med fremragende prestasjoner. Det er særlig yngre talenter lederne har brukt tid på å utvikle og hadde sett for seg at skulle bringe bedriften videre som slutter. Utfordringen forklares av en leder på følgende måte:

"Jeg tror belønningssystemet vårt påvirker turnover på sikt, fordi mange av våre 5´ere har sluttet de siste årene. (...) Det er nok en fare for at vi ikke klarer å beholde de flinkeste fordi vi ikke bruker systemet i større grad. Rammene er der, men vi er ikke kyniske nok. Når vi kalibrerer er det mer tusenlappene vi flytter på enn de store beløpene. (...) De går til konkurrenter med helt andre betingelser."

Som følge av nedgangstider i bransjen har det vært nødvendig for bedriften å kutte lønnskostnader. Karaktersystemet benyttes fortsatt, men de siste to årene har alle ansatte fått lik prosentvis lønnsøkning. Det har derfor ikke vært noen differensiering i finansiell belønning tilknyttet karakterene. Ledere forteller om dette som en utfordring, da karakterene således blir mindre betydningsfulle. Utfordringen forsterkes når det som følge av nedbemanninger i bedriften kreves mer av de ansatte, til tross for manglende finansiell

kompensasjon. Ansatte opplever derimot ikke at motivasjonen har blitt svekket i særlig grad når prestasjonsbaserte lønnstillegg har blitt fjernet. Overgangen har for dem vært liten, da det fra før var små forskjeller i belønning mellom sterke og svake ytere.

Ikke-finansiell belønning og indre motivasjon

Til tross for mangel på tilstrekkelig finansiell belønning tilknyttet karakterene, forteller ansatte at bedriften er gode på å kompensere med ikke-finansiell belønning. Flere ansatte opplever spennende arbeidsoppgaver, anerkjennelse og faglige utviklingsmuligheter like viktig som finansielle belønninger. Dette gjelder særlig de siste årene når prestasjonsbaserte lønnstillegg har blitt fjernet. Ledere har derfor prøvd å bli flinkere til å vektlegge andre faktorer enn de økonomiske for å motivere og stimulere ansatte. Videre opplyser ansatte om en innebygd yrkesstolthet, der de drives av et indre ønske om å gjøre en god jobb uavhengig av belønning. En ansatt forklarer betydningen av ikke-finansiell belønning slik:

"Jeg ser at ledelsen har blitt flinkere de siste årene til å fokusere på ikke-økonomisk belønning. Dette er jo viktig nå som den finansielle belønningen ikke er overveldende. (...) Det er helt klart at vi ønsker mer finansiell belønning i den grad det er mulig, men i disse tider er det viktigere med de rette oppgavene og bli utfordret og stimulert faglig. Jeg tror nok folk på grunn av nedgangstidene kanskje har blitt mindre opptatt av de vurderingskriteriene og belønningsmekanismene, men heller fokus på å bli ivaretatt på kompetanseutvikling og anerkjennelse."

Selv om ikke-finansiell belønning og indre motivasjon i noen grad kompenserer for små forskjeller i prestasjonsbaserte lønnstillegg, understreker ansatte at det på sikt er nødvendig å i større grad belønne gode prestasjoner. En ansatt utdyper det slik:

"Det er klart at ethvert insentivsystem også må inneholde finansielle insentiver som gjør at de som står på mest får igjen for det. Det må være differensiering i større grad enn hvordan det er nå, med tanke på hvordan man blir belønnet. Sånn som systemet er i dag, og gitt dagens marked, er det ikke rom for en slik type differensiering. Det er mye likebehandling uansett leveranse, og det er ikke bærekraftig på lang sikt. (...) Så du må ha en kombinasjon av både finansiell og ikke-finansiell belønning for at man skal være en motivert gruppe, og de må kombineres på en god måte. Før eller siden må nok bedriften bevege seg mer mot normale tilstander igjen, eller skape et nytt og bedre system."

5.1.3 Oppsummering av funn i bedrift A

Bedrift A sitt system for prestasjonsevalueringer innebærer skjønn og tvungen karakterfordeling. Skjønnsmessige evalueringer fremstår nødvendig for å fange om nyanserikdommen i ansattes arbeidshverdag. Tvungen fordeling av karakterer virker derimot å innebære utfordringer. I praksis normalfordeles ikke karakterene, og forventningen om gjennomsnittskarakter medfører lav grad av differensiering. I tillegg har bedriften relativt lave finansielle belønninger tilknyttet karakterene, noe som har medført at flere av de flinkeste ansatte har sluttet, samt lav motivasjonseffekt av prestasjonsevalueringene. Tabellen under oppsummerer de viktigste funnene i bedriften.

Oppsummerte funn i bedrift A	
Skjønnsmessige prestasjonsevalueringer	
Helhetlige vurderinger	Ja, tar hensyn til støy og kompleksitet i arbeidsoppgaver.
Observasjon av innsats	God, er flere ledere som evaluerer og kommer med innspill. Andres innspill er imidlertid mindre nyanserike.
Tilbakemeldinger	Flinke på å gi ros, men kvier seg for kritiske tilbakemeldinger.
Trynefaktor	Ansatte og ledere ser at trynefaktor kan påvirke vurderingene, men det er stort sett ikke et utstrakt problem. Dempes når det er flere som evaluerer prestasjonene.
Motvilje mot å differensiere	Ja, ledere finner det krevende å gi svake karakterer. Overvurderer ofte underytere sine prestasjoner.
Karakterer og tvungen fordeling	
Sikre faglig og personlig utvikling	Ja, men er for stort fokus på karakterer heller enn tilbakemeldinger. Karakterer og kalibrering fortrenger jevnlige tilbakemeldinger.
Transparent link mellom ytelse og belønning	Ja, karakterene innebærer omfattende dokumentasjon
Ytelsesstandarder	Kalibrering sikrer kommunikasjon mellom ledere, det gir likere ytelsesstandarder. Likevel utøves skjønn forskjellig.
Identifisere sterke, svake og normale ytere	Ja, men ansatte opplever karakterene som demotiverende merkelapper.
Differensiering	Nei, 95 % får tilsvarende karakterer. Skyldes forventning om gjennomsnittskarakter. Ledere finner det vanskelig å identifisere svake ytere, kan da heller ikke premiere gode prestasjoner. Ansatte finner det demotiverende.
Rettferdig karaktergivning	Delvis. Systemet støtter ikke teamytelse, og enkelte ledere gir karakterer etter hva som passer i fordelingen.
Karakteren 3	Skal være god, men ansatte finner den demotiverende. For mange får karakteren 3.
Målsetting	Insentiv til lite ambisiøs målsetting. Særlig for seniorer som har fått økte krav over tid.
Finansielle belønninger tilknyttet karakterene	Oppleves for små til å motivere. Har mistet ansatte som får karakteren 5, siden de ikke belønnes tilstrekkelig.
Nedgangstider	Kuttet prestasjonsbaserte lønnsøkninger.
Ikke-finansielle belønninger	Viktig, men kompenserer ikke for manglende finansielle belønninger.

Tabell 6: Oppsummering av funn i bedrift A

5.2 Funn i bedrift B

I det følgende presenteres funn fra bedrift B som er relevante for å besvare studiens problemstilling. For å sikre en systematisk gjennomgang vil vi først presentere funn som omhandler skjønnsmessige prestasjonsevalueringer, før vi presenterer funn vedrørende karakterer og tvungen fordeling.

5.2.1 Skjønnsmessige prestasjonsevalueringer i bedrift B

Bedrift B bruker skjønn i sine prestasjonsevalueringer. Skjønnsmessige vurderinger virker å innebære både muligheter og utfordringer for bedriften, og funn fra undersøkelsene vil i det følgende presenteres.

Helhetlige vurderinger

Samtlige ledere i bedrift B forteller at det er avgjørende å kunne gi helhetlige vurderinger av ansattes prestasjoner, da målene de jobber mot sjeldent er kvantifiserbare. Ansatte i bedriften har ofte komplekse arbeidsoppgaver, og en evaluering må derfor innebære at noen tolker jobben som er utført. Skjønnsmessige prestasjonsevalueringer er den eneste evalueringsformen som gir grunnlag for en helhetlig og tilpasset vurdering av den enkelte, og er derfor avgjørende for å kunne motivere til ekstra innsats gjennom prestasjonsvurderinger. En leder oppsummerer det slik:

"Hvis du skal se talentene og hva som er unikt med hver enkelt må du nesten bruke skjønn. Det går ikke an å matematisk beregne frem hva den enkelte har gjort, oppgavene er for komplekse. (...) Når man bruker skjønn får man muligheten til å ta hensyn til flere aspekter ved den enkeltes arbeidshverdag, og jeg tenker man er nødt til å vurdere helheten for å motivere. (...) Så jeg tror ikke du kommer vekk fra at du må ha en skjønnsmessig del, det vil jo være mer langsiktig positivt for alle."

Samarbeid trekkes frem av ledere som sentralt for at ansatte skal prestere godt. De anser bruk av skjønn egnet fordi det gir mulighet til å ta hensyn til samarbeidsevner i vurderingene. Videre påvirkes resultatene til de ansatte av forhold utenfor deres kontroll. Ansatte opplever det hensiktsmessig med bruk av skjønn i prestasjonsevalueringene, da det gir mulighet til å premiere god innsats selv når det ikke reflekteres av resultatene.

Observasjon av ansattes innsats

Ansatte forteller at hvorvidt prestasjonsevalueringene oppleves rettferdige avhenger av leders innsikt i deres prestasjoner og arbeidshverdag. I utgangspunktet er det nærmeste leder som foretar evaluering av sin gruppe. De ansatte synes dette fungerer godt, fordi nærmeste leder ofte har god mulighet til å observere deres innsats og prestasjoner. Samtidig er bedriften relativt liten i Norge, noe som gjør ansattes prestasjoner mer synlige.

På en annen side forteller ledere om enkelte tilfeller der det er mer krevende å observere ansattes prestasjoner, særlig om det er mangel på fysisk eller faglig nærhet. I tilfeller med fysisk eller faglig distanse opplever ansatte at hvorvidt de blir sett avhenger av egen evne til å fremme sine prestasjoner. Denne utfordringen uttrykkes særlig å gjelde for ansatte med dualrapportering. Flere og flere ansatte i bedriften har dualrapportering, der de rapporterer til en leder ved hovedkontoret i utlandet og en leder ved kontoret i Norge. Da er det lederen ved kontoret i utlandet som er faglig ansvarlig. Utfordringen knyttet til å bli sett for ansatte med dualrapportering utdypes slik av en ansatt:

"Hvis karakterene skal virke motiverende, er det veldig viktig at lederen har mulighet til å observere din innsats og prestasjoner. Hvis man ikke blir sett vil karakterene heller ha motsatt effekt. Jeg opplever generelt at min leder har god mulighet til å se meg, men jeg vet at det ikke er sånn for alle (...). Særlig de i bedriften som har dualrapportering har nevnt utfordringer knyttet til det å bli sett av sine ledere. Da er det gjerne den lederen med faglig ansvar som sitter i utlandet, og derfor har mindre mulighet til å observere innsatsen. Lederen i Norge har kanskje ikke nødvendig kunnskap til å forstå hva som ligger i den ansattes prestasjoner, og da vil det kanskje være litt flaks om man blir sett eller ikke. Da må de ansatte kanskje bruke litt mer tid på å fremme seg selv og sine prestasjoner hvis de skal få en god karakter. (...) I vår bonusmodell er det særlig viktig at lederen klarer å forstå ansattes prestasjoner, fordi karakteren man får har så store økonomiske konsekvenser."

Tilbakemeldinger

Forståelse og innsikt i ansattes prestasjoner påvirker som nevnt hvorvidt karakterene de mottar oppleves rettferdige. I tillegg vil slik innsikt være avgjørende for leders evne til å gi gode tilbakemeldinger. Uten nødvendig innsikt blir tilbakemeldingene i følge ansatte for generelle og derfor mindre motiverende. Særlig ansatte med dualrapportering ønsker mer konkrete tilbakemeldinger på sine prestasjoner.

Ansatte uten dualrapportering opplyser at de mottar tilstrekkelig med tilbakemeldinger tilknyttet måloppnåelse. I bedriften legger systemet opp til to formelle samtaler i året, men ledere har gjerne ukentlige møter med sine ansatte hvor det gis tilbakemeldinger. Ansatte opplever derfor sjeldent å bli overrasket i de formelle medarbeidersamtalene. En ansatt beskriver prosessen med tilbakemeldinger som følgende:

"De formelle medarbeidersamtalene har vi to ganger i året, der vi går gjennom målene man har satt opp i begynnelsen av året og hvordan de er utført. Når det er sagt, så får vi også kontinuerlige tilbakemeldinger. (...). Vi har ukentlige møter der vi oppdaterer lederen på det vi jobber med og får tilbakemeldinger. Sånn sett får jeg jo tilbakemeldinger på målene mine hver uke, fordi det jeg jobber med er jo nært knyttet til målene jeg har satt opp."

Trynefaktor

For de ansatte er det avgjørende at prestasjonsevalueringer foretas på rettferdig grunnlag. Dette utfordres når ledere bruker skjønn i prestasjonsevalueringene, da karaktergivningen kan påvirkes av relasjonen lederen har til ulike ansatte. Ansatte forteller at de opplever enkelte ledere mer påvirkelige av trynefaktor enn andre når de foretar evalueringene. De opplever særlig at trynefaktor har betydning når de vipper mellom to karakterer. En ansatt beskriver sin opplevelse av trynefaktor slik:

"Selvfølgelig vil det alltid være en sjarmfaktor inkludert i evalueringen når den er basert på skjønn. Noen har man bedre kjemi med enn andre, og de vil man gjerne ønske å premiere litt. Hvis man derimot kommer skjevt ut med lederen sin vil det kunne påvirke senere evalueringer negativt. (...) Ulempen er jo den at du fort kan bli straffet eller for mye belønnet ut fra en ren subjektiv oppfatning din overordnet har. (...) Jeg tenker denne sjarmfaktoren særlig får betydning hvis man vipper mellom to karakterer, for da er det litt lettere for leder å forsvare evalueringen."

Ansatte opplever også trynefaktor som et større problem når kriteriene de måles etter er mindre spesifikke og kvantifiserbare. Som beskrevet i del 4.2.2 evalueres både atferd og leveranse, og evalueringen tilknyttet atferd innebærer mindre målbare kriterier. Dette medfører en mer subjektiv og uhåndgripelig evaluering på atferd, og trynefaktoren oppleves derfor å få mer spillerom. Av ovennevnte grunner forteller ledere at det også er vanskeligere å evaluere atferd enn leveranser.

Motvilje mot å differensiere

En utfordring med skjønnsmessige prestasjonsevalueringer er når ledere foretar vurderinger basert på bekvemmelighet heller enn faktiske prestasjoner. Ledere forteller at de synes det er vanskelig å gi ansatte dårlige karakterer, da de frykter kritiske prestasjonsevalueringer kan ødelegge relasjonen de har til sine ansatte. I tvilstilfeller foretrekker lederne derfor å vippe ansatte opp, og der det er mulig unngår de å gi karakteren "under forventet". En leder forklarer motviljen mot å gi svake karakterer som følgende:

"Jeg synes det er vanskelig å gi ansatte svake karakterer, fordi det kan oppstå konflikter og dårlig stemning. (...) Hvis jeg for eksempel har en ansatt som har oppfylt kriterier til å kunne få både karakteren "under forventet" og "som forventet", blir summen ofte at jeg setter dem på "som forventet."

5.2.2 Karakterer og tvungen fordeling i bedrift B

Bedrift B bruker tvungen fordeling av ansattes karakterer, med fastsatte prosentandeler på hvor mange som skal få hver av de fire karakterene. I det følgende vil vi presentere funn fra intervjuene i bedriften om tvungen fordeling og karakterer.

Faglig og personlig utvikling

HR-leder forteller at formålet med å bruke karakterer i prestasjonsevalueringene av de ansatte, er å tilrettelegge for ansattes faglige og personlige utvikling. Ved å både sette mål og nøye evaluere disse som bakgrunn for karaktersettingen, gir karakterene et godt grunnlag for å følge opp den enkelte ansatte. I tillegg kan relativt få ansatte få toppkarakterer, og de har derfor stadig noe å strekke seg etter. Lederen beskriver det slik:

"Ansatte vil ved systemet vi bruker stadig ha en gulrot å jobbe etter. Om man har fått "som forventet" flere år vil man jobbe mer for å se om man klarer å få en høyere karakter. På den måten tror jeg karaktersystemet virker motiverende for de ansatte."

Tilbakemeldinger

Enkelte ledere blir som følge av kalibreringsprosessen mer forsiktige med å gi tilbakemeldinger. Vanligvis gir ledere ansatte en indikasjon på hvilken karakter de ligger an til halvveis i året, og har også da en liten kalibrering. Ledere forteller de i slike samtaler vegrer seg for å gi gode tilbakemeldinger. Dette skyldes at de ikke ønsker å gi falske forhåpninger i tilfelle kalibreringen resulterer i lavere karakterer enn de så for seg. Lederne

sier dette er særlig utfordrende fordi det er så store forskjeller i finansielle belønninger tilknyttet de ulike karakterene, og det oppleves derfor ekstra demotiverende for ansatte som får lavere karakter enn de forventet.

Ytelsesstandarder og identifisering

Formålet med å kalibrere ansattes karakterer er å skape felles forståelse for hvordan karakterskalaen brukes, og rettferdig karaktergivning uavhengig av hvilken leder man har. Ansatte opplever det viktig med lik bruk av karakterskalaen i de ulike avdelingene, særlig fordi karakterene er tilknyttet vesentlige finansielle belønninger. Imidlertid forteller ledere om forskjeller i hvordan skjønn utøves, noe som medfører at enkelte ledere lettere gir ut gode karakterer enn andre. En leder forteller om betydningen av like ytelsesstandarder på følgende måte:

"Å ha en kalibrering her på huset er viktig for å sikre delt forståelse for hvordan karakterene og karakterskalaen brukes. Det skal ikke være sånn at man foretrekker å jobbe for en leder siden han lettere gir ut de gode karakterene."

I bedriften foreligger det klare retningslinjer for hvordan karakterer skal settes. Dette er ment å lette prosessen med karakterfastsettelse for lederne. Imidlertid forteller en leder om utfordringer ved å bruke retningslinjene, fordi ansatte ofte tilfredsstiller krav fra ulike kategorier. Det oppleves særlig utfordrende å skille mellom karakterene "som forventet" og "over forventet".

Forhåndsbestemte prosentandeler

Flere ledere forteller om utfordringer når ansattes karakterer må ligge innenfor en forhåndsbestemt fordeling. Særlig oppleves det utfordrende når karakterene i små grupper skal fordeles. Lederne synes ansatte leverer godt, og opplever det vanskelig å bestemme hvem som skal få de ulike karakterene. Det er klare retningslinjer på hvor mange som kan få de ulike karakterene, men lederne opplever at det alltid er flere enn tillatt som burde få de beste karakterene. Det er særlig vanskelig å skille ansatte som ligger på grensen mellom "som forventet" og "over forventet". Hvem som da ender opp med "over forventet" vil avgjøres av den relative rangeringen, men også av tilfeldigheter. HR-leder forklarer nevnte utfordring med forhåndsbestemt fordeling av ansattes karakterer på følgende måte:

"Jeg har vært med i prosesser hvor man har sittet med fordelingskurven og vi har sett at vi må jo ha noen med de dårligere karakterene. Det er kjedelig når noen blir plukket ut fordi de

ikke nødvendigvis var mindre gode enn de andre, men fordi vi nettopp må gi noen de dårligere karakterene. (...) Når vi sammenligner to personer med relativt like prestasjoner vil vi ofte tenke på hvem vi flyttet opp i fjor, så tenker vi kanskje at den andre burde bli flyttet opp i år. Men det er litt leit, for vi ønsker jo å premiere folk. Når to personer har omtrent like prestasjoner føles det hardt å måtte sette den ene ned utelukkende fordi systemet tvinger oss til det."

En annen effekt av bedriftens forhåndsbestemte fordeling er at det ofte er vanskelig å få bedre karakter enn "som forventet", selv om man øker innsatsen. Ansatte synes dette reduserer motivasjonseffekten av karakterene. Ved øvre tak på antall som kan få høyere karakterer, blir det vanskelig å premiere forbedrede prestasjoner. Karakteren er gjerne det som fester seg hos de ansatte, og da er det ikke tilstrekkelig om leder gir uttrykk for å legge merke til forbedringene hvis det ikke synes i karakteren som blir gitt. En leder forklarer det slik:

"Vi har en utfordring med de som år etter år havner i "som forventet", og ikke kommer seg opp på "over forventet" fordi det er noen som har gjort enda mer. De er ofte veldig flinke og har jobbet hardt i alle år, men uansett hva de gjør forblir de i denne kategorien på grunn av den relative rangeringen.(...) Det er veldig leit fordi det er utrolig mange flinke folk som vi gjerne skulle gitt "over forventet", men systemet vårt tillater ikke det."

Differensiering

Bedriften bruker som nevnt en form for tvungen fordeling som innebærer forhåndsbestemte krav til andel av ansatte som kan få hver karakter. Ledere har derfor ikke mulighet til å gi en annen karakterfordeling enn det som er bestemt fra sentralt hold. Slik sikrer bedriften spredning i karakterer, noe de ansatte ser på som viktig om systemet skal virke motiverende. I tillegg forteller ledere at tvungen fordeling av ansattes karakterer gjør dem mindre motvillige mot å gi ansatte svake karakterer. Da kalibreringsprosessen er etter medarbeidersamtalen, kan ledere være mindre ærlige når de gir karakterantydningen til sine ansatte i samtalen. Når den endelige karakteren gis kan nærmeste leder skylde på kalibreringen fremfor å ta ansvar for vurderingen. En leder utdyper som følgende:

"Når jeg må gi ansatte karakteren "under forventet" synes jeg alltid det er ubehagelig. Noe som likevel gjør det enklere er at kalibreringsprosessen er etter jeg gir karakterer til mine ansatte. Et eksempel kan være hvis jeg har en ansatt jeg egentlig vurderer som "under

forventet", men tilfredsstiller enkelte kriterier på "som forventet". Da kan jeg i medarbeidersamtalen si at jeg vurderer dem til et sted mellom "som forventet" og "under forventet", men at kalibreringen vil avgjøre hvilken karakter som blir den endelige. Når den endelige karakteren da blir "under forventet" er det ikke bare jeg som bestemte at det ble sånn, og da er det enklere å opprettholde relasjonene."

Karakteren "som forventet"

Karakteren "som forventet" er ment å være en god karakter, og der 70-80 % skal ligge i prestasjonsevalueringene. Flere ledere og ansatte har uttrykt misnøye når så mange mottar samme karakter. Gruppen omfatter ansatte som har den respektive karakteren på både atferd og leveranse, samt ansatte som har enten "over forventet" eller "under forventet" på et av områdene. Karakteren oppfattes derfor å innebære et for bredt spekter av prestasjoner og individer. Både ledere og ansatte har uttrykt et ønske om å innføre en bredere karakterskala. En ansatt illustrerer utfordringen gjennom følgende sitat:

"Det er et stort spenn av prestasjoner innenfor karakteren "som forventet". Hvis du hadde laget et nummerert skala fra 1-10, ville kanskje 2-9 ligget innenfor denne karakteren. (..) Det kunne kanskje vært flere kategorier på midten, i stedet for å bare ha den ene."

Utfordringen blir forsterket fordi karakterskalaen med fire kategorier er asymmetrisk. Det er vanskelig å skaffe forståelse for at "som forventet" er en god karakter, når den er nest nederst på skalaen. Om en ansatt får karakteren "under forventet" havner man på en forbedringsliste og får tett oppfølging, og det skal mye til for å havne i denne kategorien. Ansatte opplever derfor ikke karakteren rett over denne som god, og blir ofte skuffet når det er "som forventet" de mottar. Dette gjelder særlig når det er to karakterer som er bedre enn "som forventet", men kun en som er svakere. En ansatt forteller om karakteren "som forventet" slik:

"Det er litt rart at når du har fire kategorier å velge mellom, så er selskapet fornøyd med den nest dårligste. Den settingen er vanskelig å forstå. (...) Sånn jeg har skjønt det skal kun et fåtall havne på "under forventet", og de som gjør det skal få beskjed på forhånd siden det er en så hard karakter å få. Da blir det vanskelig å forstå at karakteren over det betyr at man gjør en god og tilfredsstillende jobb. (...) Jeg ser ikke hvordan du skal bli supermotivert av å få en 2'er når du kunne fått en 3'er eller 4'er."

Finansiell belønning

Bedriften har et generøst bonussystem, noe de håper virker motiverende på de ansatte. Samtidig er det store forskjeller i bonusutbetalingene, som innebærer at systemet for mange kan virke urettferdig. Dette gjelder spesielt når ansatte med tilsvarende prestasjoner kan ende med svært ulike økonomiske belønninger. Et eksempel kan være når ledere har to ansatte de ønsker å gi "over forventet", men de kun har plass til en i fordelingen. Det kan da være tilfeldigheter som avgjør hvem av de to som ender med karakteren. Forskjellen mellom de to ansattes belønning kan i slike tilfeller utgjøre mange hundre tusen. Den ansatte som blir vippet ned i karaktersettingen kan ende opp demotivert i stedet for å verdsette faktisk mottatt bonus. Systemet som i utgangspunktet er gunstig for ansatte, kan oppleves urettferdig når forskjellene blir så store. En leder utdyper det slik:

"For ansatte som vipper mellom to karakterer har det vanvittig mye å si for bonusen hvilken karakter man ender med å få, det kan utgjøre mange hundre tusener for den enkelte. For oss som avgjør blir det noen diskusjoner, så ender det med at en blir trukket opp en karakter og en blir trukket ned. (...) Jeg som leder ønsker jo å premiere folk, men det føles hardt når forskjellene blir så store økonomisk. De ansatte er kanskje nesten like gode, men så ender den ene opp med flere hundre tusen mer enn den andre. Den som fikk den laveste bonusen glemmer da at den faktisk har fått en god bonus, for den sammenligner seg med den andre. Et system som skal virke motiverende blir på den måten ganske demotiverende."

Til tross for utfordringer knyttet til store forskjeller i finansielle belønninger, forteller flere ansatte om bonusene som en driver for motivasjon. Ansatte anser seg godt ivaretatt økonomisk, grunnet konkurransedyktig fastlønn kombinert med generøse bonusordninger. De ser på forskjeller i finansielle belønninger tilknyttet karakterene som hensiktsmessig, da de som legger ned ekstra mye innsats faktisk får uttelling for det. En ansatt forklarer sine oppfatninger i følgende sitat:

"Belønningssystemet vi har er en god motivasjonsdriver. Fordi vi allerede har en solid grunnlønn vil man uansett om man ender med minstebonusen ha det godt økonomisk. Jeg tenker at vi som jobber her er fornøyde, og tenker at vi er heldige, og det reflekteres jo i den jobben vi gjør(...). Det er fint når man bruker karakterer at det gis skikkelig belønning når man gjør det bra. Hvis ikke hadde det kanskje vært litt meningsløst å bruke så mye tid på å gi disse karakterene."

Både ledere og ansatte hevder store finansielle belønninger virker positivt på formålet om å tiltrekke og beholde ønsket arbeidskraft. Flere ansatte forteller at de har valgt å starte i bedriften nettopp grunnet god økonomisk kompensasjon. I tillegg ser de på belønningene som motivasjon til å bli værende i bedriften. Ansatte opplever betingelsene som bedre enn hva andre operatørbedrifter i bransjen tilbyr. En ansatt oppsummerer effekten på turnover slik:

"Jeg tror belønningssystemet vårt har en effekt både på at folk ønsker å starte her og at de ønsker å bli. Jeg vet våre betingelser er bedre enn hos mange andre konkurrerende selskaper. Selv i gode tider da mulighetene var bedre for å bytte jobb, var det ingen som sluttet så vidt jeg vet. Det må jo bety at belønningssystemet her fungerer ganske godt."

Bedriften har valgt nulloppgjør i lønnsøkninger som følge av nedgangstidene i bransjen de siste årene. Likevel har de valgt å beholde prestasjonsbaserte individuelle bonuser, og bonusutbetalingene har vært av tilsvarende størrelse som i et normalår. Ansatte som fikk karakteren "som forventet" mottok i fjor 85 % av bonusen de ville fått i normalår, mens ansatte med karakteren "over forventet" mottok 95 % av det de ville fått i normalår. I følge ledere har det har vært viktig å belønne de som er igjen i bedriften, fordi nedbemanninger har ført til at det kreves mer av hver enkelt. Både ansatte og ledere har opplevd bevaring av bonusene svært motiverende, da de til tross for nedgangstider har blitt belønnet for ekstra innsats og gode prestasjoner.

Ikke-finansiell belønning

Ansatte opplever bonusen de mottar som veiledende, og som en bekreftelse på gode prestasjoner. Når lønnen ansatte har er konkurransedyktig, anser de imidlertid andre faktorer enn finansielle som viktige for motivasjonen i det daglige. Samtlige ansatte i bedriften forteller at ikke-finansiell belønning som spennende arbeidsoppgaver, anerkjennelse og utviklingsmuligheter, er viktig for deres motivasjon. En ansatt forklarer sine perspektiver på følgende måte:

"Bonusen veileder meg til å gjøre de riktige tingene på den riktige måten, og gir meg et ønske om å bli værende i bedriften. (...) Det er klart at bonusen har betydning for motivasjonen, men når man kommer over et visst lønnsnivå får bonusen mindre betydning. Da vil andre faktorer som anerkjennelse i form av nye og interessante oppgaver og muligheter for utvikling være det viktigste."

5.2.3 Oppsummering av funn i bedrift B

Bedrift B sitt system for prestasjonsevalueringer innebærer skjønn og tvungen karakterfordeling. Skjønnsmessige evalueringer fremstår nødvendig for å fange om nyanserikdommen i ansattes arbeidshverdag, men ansatte oppfatter trynefaktor som et utstrakt problem. Tvungen fordeling av karakterer sikrer differensiering i karakterer, og identifisering av sterke, normale og svake ytere. Imidlertid medfører tvungen fordeling i enkelte tilfeller urettferdig karaktergivning basert på hva som passer i fordelingen, heller enn karakterer som reflekterer ansattes faktiske prestasjoner. Videre har bedriften rause finansielle belønninger tilknyttet karakterene. Generøse bonusordninger bidrar til å tiltrekke og beholde ønsket arbeidskraft, og til å gjøre karakterene mer motiverende. Tabellen under oppsummerer de viktigste funnene i bedriften.

Oppsummerte funn i bedrift B	
Skjønnsmessige prestasjonsevalueringer	
Helhetlige vurderinger	Ja, tar hensyn til støy og kompleksitet i arbeidsoppgaver.
Observasjon av innsats	Passe god. Kun en leder som evaluerer ansattes prestasjoner. Leder har som regel god innsikt, men i tilfeller det ikke er gjeldende avhenger observasjon av ansattes evne til å fremme egne prestasjoner.
Tilbakemeldinger	Flinke til å gi ros, men kvier seg for kritiske tilbakemeldinger.
Trynefaktor	Ansatte og ledere ser at trynefaktor kan påvirke prestasjonsevalueringene. Krevende at det kun er en som evaluerer de ansatte. Oppleves problematisk fordi det har store økonomiske konsekvenser hvilken karakter man får.
Motvilje mot å differensiere	Ja, ledere finner det krevende å gi svake karakterer. Problemet dempes av kalibrering etter medarbeidersamtalen og fastsatte prosentandeler som skal få hver karakter.
Karakterer og tvungen fordeling	
Sikre faglig og personlig utvikling	Ja, men er for stort fokus på karakterer heller enn tilbakemeldinger. Karakterer og kalibrering fortrenger jevnlige tilbakemeldinger.
Transparent link mellom ytelse og belønning	Ja, karakterene innebærer omfattende dokumentasjon.
Ytelsesstandarder	Kalibrering sikrer kommunikasjon mellom ledere, det gir likere ytelsesstandarder. Likevel utøves skjønn forskjellig.
Identifisere sterke, svake og normale ytere	Ja, nevnes ikke problem med at karakterer oppleves som demotiverende ”merkelapper”.
Differensiering	Ja, sikres gjennom fastsatte prosentandeler som kan få hver karakter. Ledere kan ikke avstå fra å differensiere.
Rettferdig karaktergivning	Delvis. Systemet støtter ikke teamytelse. Tilfeldigheter avgjør når flere enn tillatt fortjener samme karakter. Ansatte med tilsvarende prestasjoner kan ende med ulik karakter.
Karakteren ”som forventet”	Skal være god, men ansatte finner den demotiverende. For mange får karakteren ”som forventet”.
Målsetting	Insentiv til lite ambisiøs målsetting.
Finansielle belønninger tilknyttet karakterene	Oppleves rause og motiverer. Belønningene bidrar til å tiltrekke og beholde ønsket arbeidskraft.
Nedgangstider	Kuttet prestasjonsbaserte lønnsøkninger, men har beholdt bonuser som er av tilsvarende størrelse som i normalår.
Ikke-finansielle belønninger	Viktig komplement til store finansielle belønninger.

Tabell 7: Oppsummering av funn i bedrift B

5.3 Funn i bedrift C

Vi vil nå presentere funn fra bedrift C som er relevante for å besvare studiens problemstilling. For å sikre en systematisk gjennomgang vil vi først presentere funn som omhandler skjønnsmessige prestasjonsevalueringer, før vi presenterer funn vedrørende karakterer og tvungen fordeling.

5.3.1 Skjønnsmessige prestasjonsevalueringer i bedrift C

I likhet med bedrift A og B, bruker bedrift C bruker skjønn i prestasjonsevalueringer. Skjønnsmessige vurderinger virker å innebære både muligheter og utfordringer for bedriften, og funn fra intervjuene presenteres i det følgende.

Helhetlige vurderinger

I bedrift C forteller ansatte at deres arbeidsoppgaver er sammensatte og med mål som er vanskelige å kvantifisere. Det er derfor nødvendig med skjønnsmessige prestasjonsevalueringer om vurderingen skal være dekkende for deres arbeidshverdag. Ved vurderinger basert på skjønn opplyser ledere om mulighet for å ta hensyn til aspekter det er vanskelig å måle objektivt, men som er viktig for å fange opp den enkeltes bidrag til verdiskapningen. En leder forklarer hensikten med å benytte skjønn slik:

"En skjønnsmessig vurdering er nødvendig for å kunne gi best mulig tilbakemeldinger på ansattes prestasjoner. Det er ikke mulig å måle dem ved hjelp av objektive mål, fordi de ansatte har mange og sammensatte arbeidsoppgaver. (...) Hos oss er veien til mål like viktig som hvilke resultater man oppnår. For eksempel er hvor godt de ansatte samarbeider og hvor mye stå-på-vilje de har viktig å vurdere. Slike nyanser fanges ikke opp med kvantitative mål (...) For å klare å motivere ansatte på riktige områder gjennom evalueringene må vi se helheten i det de leverer, og da er vi nødt til å bruke skjønn."

Skjønnsmessige vurderinger er også hensiktsmessig når ansattes prestasjoner til dels avhenger av ukontrollerbare forhold. En ansatt forteller at det er hvordan du løser arbeidsoppgavene som blir vurdert av lederen, og dersom man møter på utfordringer som medfører svakere resultater vil dette tas hensyn til i prestasjonsevalueringen. Den ansatte forklarer dette som følgende:

"Fordelen ved at ledere vurderer oss skjønnsmessig er at vi kun blir målt på områder vi kan kontrollere med egen innsats. Det kan være mange årsaker til at du for eksempel ikke har nådd alle målene dine, og evalueringene de gjør kan da ta hensyn til disse. Det er prosessen hvordan du løser oppgaven som blir grunnlaget for om du får en god eller dårlig score, ikke om du har flaks eller uflaks. (...) Jeg opplever derfor at bruken av skjønn er hensiktsmessig."

Observasjon av ansattes innsats

Ansatte forteller at hvorvidt prestasjonsevalueringene oppleves rettferdige avhenger av leders innsikt i deres arbeidshverdag og prestasjoner. I bedrift C evalueres hver ansatt av linjeleder og lederen på nivået over linjeleder. Dersom en ansatt har en rolle som går ut mot mange andre i bedriften skal også deres innspill tas hensyn til i vurderingen. Gruppene den enkelte leder skal evaluere er relativt små, og de har derfor god kjennskap til de ansattes prestasjoner. Ansatte opplever på sin side lederes mulighet til å observere deres innsats og prestasjoner som god. Imidlertid forklarer ansatte at hvorvidt man blir godt observert avhenger av om man har en god leder, og avstanden mellom leder og ansatt. En ansatt sier det på følgende måte:

"Jeg opplever generelt at lederne her har god oversikt over ansatte sin arbeidshverdag.(...) Det er ganske små forhold her, så ledere har god kjennskap til ansattes innsats og prestasjoner. Men det varierer selvfølgelig litt, for noen ledere er mye bedre på å sette seg inn hva man jobber med enn andre.(...) At jeg føler meg sett er avgjørende hvis jeg skal oppleve karakteren som rettferdig."

En ansatt forklarer at dersom man ikke føler seg riktig oppfattet eller observert kan man klage og be om vurdering fra andre enn linjelederen sin. Klagemuligheten oppleves som en sikkerhetsventil mot svake observasjoner, da ansatte alltid har mulighet til å be om vurdering fra andre de mener har viktig innsikt i deres prestasjoner.

Tilbakemeldinger

Som nevnt påvirker leders innsikt i ansattes arbeidshverdag hvorvidt prestasjonsevalueringene oppleves rettferdige. En ansatt forteller at slik innsikt også er avgjørende for leders evne til å gi motiverende tilbakemeldinger. Konstruktive tilbakemeldinger er viktig om ansatte skal ha muligheten til å bedre sine prestasjoner, og ros er nødvendig om ansatte som presterer godt skal fortsette med samme engasjement. En leder

forteller at de derfor ønsker å gi tilbakemeldinger som er konkrete og umiddelbare, og at medarbeidersamtalen skal være en oppsummering av årets tilbakemeldinger.

HR-leder beskriver jevnlige tilbakemeldinger som særlig viktige når ansatte presterer under forventet. Da det oppleves vanskelig å få en dårlig karakter, er det spesielt viktig at det ikke kommer som en overraskelse i medarbeidersamtalen. I tillegg gir ledere gjennom ærlige tilbakemeldinger ansatte mulighet til å forbedre seg om de observerer svakere prestasjoner. Imidlertid er det de ærlige og kritiske tilbakemeldingene som oppleves mest utfordrende for ledere å gi.

Ansatte forteller at lederne stort sett gir jevnlige og nyttige tilbakemeldinger. De synes lederne både gir ros når de er fornøyde med innsatsen, og kritiske tilbakemeldinger når noe må endres. På en annen side opplever ansatte forskjeller mellom ledere, der enkelte er mer ærlige og tydelige i samtaler med sine ansatte enn andre. Ansatte forteller om mangel på tilbakemeldinger som en kilde til konflikt i prestasjonsevalueringene. En ansatt beskriver sine opplevelser rundt tilbakemeldinger som følgende:

"Jeg synes det er veldig viktig at leder som gir tilbakemeldinger faktisk har god innsikt i det man jobber med, hvis ikke vil ikke tilbakemeldingene treffe. (...) Det er veldig viktig med tilbakemeldinger når man har gjort noe om er bra, for det virker motiverende. Men jeg synes kanskje det er enda viktigere med tilbakemeldinger når jeg ikke presterer etter forventningene, hvis ikke får jeg ikke mulighet til å vite hvordan jeg kan forbedre meg. Det er kanskje her det er størst forskjell mellom ledere, for noen er veldig gode på å være oppriktige, mens andre er mer forsiktige. (...) Men jeg synes generelt vi har en god kultur for tilbakemeldinger her."

Trynefaktor

Skjønnsmessige prestasjonsevalueringer kan både påvirke og bli påvirket av relasjonen mellom leder og ansatt. Ansatte forteller at enkelte ledere lettere lar seg påvirke av hvem de bevisst eller ubevisst liker bedre når de fastsetter karakterer. Både ansatte og ledere beskriver likevel trynefaktor som lite problematisk i bedriften. Dette knyttes til at det ikke kun er linjeleder som foretar prestasjonsevalueringen. Om en ansatt har en vanskelig relasjon til sin linjeleder, kan leder på nivået over redusere trynefaktorens spillerom. I tillegg forteller ansatte om lav terskel for å klage og be om vurdering fra andre enn linjeleder hvis de føler seg urettferdig behandlet.

Motvilje mot å differensiere

Ansatte forteller at de som yter mer enn forventet må få bedre karakterer enn ansatte med svakere prestasjoner om karakterene skal oppleves motiverende. Sånn det er nå får brorparten av ansatte karakteren 3. Ledere forteller om motvilje mot å bruke hele karakterskalaen, fordi de finner det ubehagelig å gi dårlige karakterer. Dette knyttes til at de frykter det kan påvirke relasjonen til ansatte negativt. Motviljen mot å gi svake karakterer er i henhold til HR-leder synlig i karakterfordelingen, da det er færre karakterer under 3 enn over 3. En leder utdyper motviljen mot å gi svakere karakterer slik:

"Min erfaring er at ledere kvier seg for å gi svake karakterer til de ansatte. (...) Jeg vet om situasjoner der ledere har klaget på enkelte ansattes prestasjoner, men likevel gitt dem karakteren 3 for å unngå misnøye og ubehagelige situasjoner. (...) Jeg tror det er med tanke på at effekten av å gi en dårlig karakter vil være mer tung og langvarig enn den positive motivasjonseffekten av å gi en god karakter.(...) Jeg vet med meg selv at skal jeg gi noen i min gruppe karakteren 2 må jeg gi noen 4. Når jeg vet at den 4'eren vil ha så dårlig effekt på den ene ansatte gir jeg heller alle 3."

5.3.2 Karakterer og tvungen fordeling i bedrift C

Bedrift C bruker tvungen fordeling av ansattes karakterer. De har en forventning om normalfordeling og et krav om gjennomsnittskarakter på 3, noe som innebærer både muligheter og utfordringer. I det følgende vil vi presentere funn fra intervjuene i bedriften vedrørende karakterer og tvungen fordeling.

Faglig og personlig utvikling

HR-leder forteller at formålet med å bruke karakterer er å tilrettelegge for ansattes faglige og personlige utvikling. Ved å både sette mål, og nøye evaluere disse som bakgrunn for karaktersettingen, gir bruk av karakterer et godt grunnlag for å følge opp den enkelte ansatte. Videre krever bruk av karakterer omfattende dokumentasjon på hvordan den enkelte er evaluert, noe som medfører et mer transparent forhold mellom evalueringer og finansielle belønninger.

Imidlertid uttrykker HR-leder et ønske om å gå bort fra bruk av karakterer, fordi lederes oppmerksomhet i for stor grad rettes mot fastsettelse av karakterene. Dette anses uheldig, da det sentrale burde være å skaffe innsikt i hvordan den enkelte medarbeider kan forbedre sine prestasjoner. I tillegg blir ansattes oppmerksomhet rettet mot karakterene heller enn

tilbakemeldingene de får samtidig i medarbeidersamtalen. Samlet medfører bruk av karakterer et for ensidig fokus hos både ledere og ansatte. HR-leder forklarer det slik:

"*Jeg synes karaktersystemet gir et litt for sterkt fokus på scoren i seg selv, og det krever så mye tid og ressurser å sette karakterer. Jeg synes det er mye bedre å gå konkret inn på de tingene som ligger til grunn for karakteren enn selve scoren.(...) Det hadde vært bedre å bare ha målsettinger for året, og gå gjennom og gi tilbakemeldinger på disse i medarbeidersamtalen. Da hadde det vært enklere å gi direkte tilbakemeldinger på områder ansatte må forbedre seg. (...) Jeg er mye mer tilhenger av den jevnlige tilbakemeldingen.*"

I følge en leder er en annen utfordring med bruk av karakterer at en svak karakter gir en mer negativ motivasjonseffekt enn en god karakter gir positiv motivasjonseffekt. Lederen beskriver utfordringen med å gi svake karakterer på følgende måte:

"En god score kan gi deg en liten motivasjonsboost der og da, men en dårlig score gjør så mye mer. Hvis jeg hadde fått en god karakter hadde det hatt en positiv effekt på meg i to uker, men det er en kortsiktig gevinst. Hadde jeg derimot fått en dårlig karakter, altså ikke hadde levd opp til leders forventninger, hadde det gjort vondt i lang tid. Nedsiden er så mye større enn oppsiden ved å bruke karakterer synes jeg."

Ledere finner det også uheldig at systemet ikke gir insentiv til ambisiøs målsetting, når det er ment å bidra til faglig og personlig utvikling. Fordi karakteren man får avhenger av hvor godt man når fastsatte mål, gis det insentiv til å sette enklere mål i begynnelsen av året.

Ytelsesstandarder

Bedriften har klare retningslinjer for hvordan karakterer skal fastsettes for å gjøre det enklere for ledere å identifisere hvem som skal få hvilken karakter. I tillegg kalibreres karakterene for å sørge for konsistens i karaktergivning på tvers av forretningsenheter. På den måten skal kalibreringen fungere som en kvalitetssikring for å fremme rettferdig karaktersetting uavhengig av hvilken leder man har. Imidlertid forteller ledere at enkelte lettere gir ut gode karakterer enn andre, noe som utdypes slik:

"Generelt er jeg fornøyd med sånn skjønnet brukes her, fordi det er så klare retningslinjer for hvordan vi skal sette karakterer. Det er en ganske teknisk oppbygning av systemet, der målet er normalfordeling av karakterene vi gir de ansatte. (...) Likevel ser jeg noen avdelinger der nesten alle får toppkarakterer, mens andre avdelinger har en mer spredning i

karakterer. Det er ikke realistisk at alle i en avdeling yter mer enn det som forventes, og det tyder på at lederen lettere gir ut gode karakterer enn andre. (...) Ulike ledere bruker nok skjønnet ulikt til å evaluere sine ansatte."

Differensiering

Alle ledere er fra sentralt hold gitt beskjed om å bruke hele karakterskalaen, og karakterene skal følge en normalfordeling. Dette er ønskelig for å sikre en spredning i karakterer. Likevel forteller både ansatte og ledere om lite bruk av karakterene 1 og 5, og at flertallet av de ansatte mottar karakteren 3. Med forventning om gjennomsnittskarakter på 3, har ledere ikke mulighet til å gi ansatte gode karakterer hvis ikke andre samtidig får dårligere karakterer. Ledere opplever sine ansatte som kunnskapsrike, og de finner det derfor krevende å identifisere hvem som skal få karakterene under 3. Med forventning om gjennomsnittskarakter på 3, blir det da også vanskelig å premiere gode prestasjoner. Ledere tilskriver derfor lav grad av differensiering til forventningen om gjennomsnittskarakter.

Ledere opplever forventning om normalfordeling og snitt på 3 egnet i større grupper og i bedriften som helhet. Dette fordi det sikrer spredning i karakterer og tilstrekkelig differensiering mellom gode og svake prestasjoner. På den andre siden uttrykker de misnøye når dette også gjelder for mindre grupper, fordi det i små grupper er mer urealistisk at ansattes prestasjoner passer inn i en gitt fordeling. En leder utdyper det i følgende sitat:

"For ledere av mindre grupper innebærer systemet at dersom man skal premiere en ansatt med karakteren 2, må en annen få karakteren 4. (...) Et eksempel på når dette er problematisk er hvis alle i gruppen leverer som forventet unntatt en som har levert bedre. Man kan da ikke premiere den beste ansatte hvis ikke en annen får en svakere karakter enn fortjent.(...). Det er sjeldent normalfordeling eller gjennomsnittskarakter passer i små avdelinger, så i de avdelingene ser man ofte at det skilles mindre mellom ansatte i evalueringene."

En annen leder forteller om situasjoner den har valgt å premiere ansatte uten å sette ned andre i gruppen. Det er altså mulig for ledere å unngå fordeling av sine ansatte i spesielle tilfeller der flere har prestert signifikant over eller under forventning.

På bakgrunn av lav grad av differensiering, ytrer både ledere og ansatte et ønske om bredere karakterskala. I følge ansatte er differensiering i karakterer avgjørende om systemet skal ha

en motiverende effekt. Når de fleste mottar samme karakter oppleves karaktersystemet mindre hensiktsmessig, fordi det ikke gir et nyansert bilde av ansattes prestasjoner.

Finansiell belønning

I bedrift C er prestasjonsevalueringene, som nevnt i del 4.2.3, kun knyttet til lønnsøkninger og eventuell tilgang til aksjeprogrammet, med mindre man er på ledernivå og har tilgang til individuell bonus. En ansatt forklarer at individuelle lønnsøkninger ikke oppfattes som en driver for motivasjon, da de er for små til å ha en signifikant effekt på personlig økonomi. I tillegg til generelt lave lønnsøkninger, er det liten finansiell forskjell knyttet til hvilken karakter man får. Ansatte forteller det er manglende gevinst ved ekstraordinære prestasjoner og å få de beste karakterene. I henhold til lederne er de økonomiske rammene for å dele ut belønninger små, noe som er bakgrunnen for små forskjeller i belønning. Utfordringen med lave prestasjonsbaserte lønnsøkninger har blitt ytterligere forsterket som følge av nedgangstidene i bransjen, med tilhørende nulloppgjør i lønnsøkninger. Samlet gir mangel på differensiering i belønning lavere motivasjonseffekt av karakterene, noe HR-lederen utdyper slik:

"Jeg tror ikke belønningssystemet vårt har noen særlig effekt på motivasjonen. Karakteren man får, og påvirkningen det har på lønnen, er ikke stor nok til at det betyr noe for de ansatte. (...) Systemet og metodikken er fin for å skille mellom gode og dårlige prestasjoner, men med små rammer blir det også mindre mulighet til å differensiere i belønning. (...) Jeg synes det er litt trist når folk har gitt en skikkelig innsats uten at det blir verdsatt nok. Her får man gjerne ikke så stor avkastning på innsatsen dessverre, og det kan oppleves urettferdig for de som yter mest."

Til tross for generelt lave forskjeller i prestasjonsbaserte lønnsjusteringer, oppleves det motiverende med bedriftens prestasjonsbaserte aksjer og spontane "eksepsjonelle belønninger". Særlig fremheves "eksepsjonelle belønninger", fordi de gir mulighet til å premiere fremragende prestasjoner umiddelbart. En ansatt forteller at "eksepsjonelle belønninger" er motiverende, og mer enn kompenserer for små forskjeller i lønnsøkninger, særlig fordi de er uforutsette. Den ansatte forklarer det slik:

"Hvis jeg har gjort en ekstraordinær innsats har det skjedd flere ganger at ledelsen har kommet med en uforutsett bonus som jeg får der og da. (...) Når den kommer overraskende

og jeg ikke har forventet bonusen eller gått og ventet på den, er den veldig motiverende. Mye mer enn ethvert lønnstillegg jeg kunne fått."

Flere ansatte anser også mindre forskjeller i lønnsøkninger som en positiv side med bedriftens belønningssystem. De mener det gir en følelse av at samtlige ansatte jobber kollektivt for fellesskapets beste, og at alle ansattes bidrag er like viktig. En ansatt forklarer dette slik:

"Det er greit med litt forskjell i individuell belønning, bare det ikke blir for stor forskjell. Du skal merke at du er et lokomotiv og en som drar, men det må ikke være sånn at du får en ekstremt mye høyere lønn enn de andre, for da blir det for synlig til slutt. Man kan gi en ekstra påskjønnelse, men det skal ikke være for stor lønnsforskjell mellom to i samme avdeling, det ville blitt helt feil og føltes veldig urettferdig. Alle ansattes bidrag er verdifulle, og det må også vises gjennom belønningssystemet."

Bedriften ønsker at belønningssystemet skal bygge opp under formålet med å tiltrekke og beholde ønsket arbeidskraft. HR-leder forteller om lav grad av turnover, også i gode tider. Imidlertid er ikke bedriften markedsledende på lønn, og enkelte ansatte har derfor gått over til andre konkurrenter med bedre betingelser. I følge ledere er det ikke typisk de flinkeste som slutter, da de er godt ivaretatt gjennom "eksepsjonelle belønninger" og prestasjonsbaserte aksjer. De som gjerne slutter er heller ansatte som presterer over forventet, men ikke får toppkarakter. Dette skyldes at deres belønninger er tilsvarende som ansatte med mindre arbeidsinnsats sine belønninger.

Ikke-finansiell belønning

Samtlige ansatte i bedriften forteller om finansielle belønninger som viktig i ethvert insentivsystem, men at ikke-finansielle prestasjonsbaserte belønninger ofte er vel så viktig. Det er særlig spennende arbeidsoppgaver, utviklingsmuligheter og tilbakemeldinger som blir trukket frem. Videre forteller ansatte at yrkesstolthet er en viktig driver for deres motivasjon til å yte høy innsats.

5.3.3 Oppsummering av funn i bedrift C

Bedrift C sitt system for prestasjonsevalueringer innebærer skjønn og tvungen karakterfordeling. Skjønnsmessige evalueringer fremstår nødvendig for å fange om nyanserikdommen i ansattes arbeidshverdag. Klagemuligheten, og at flere ledere evaluerer hver enkelt ansatt, motvirker ansattes opplevelser av urettferdige prestasjonsevalueringer. Tvungen fordeling av karakterer virker derimot å innebære utfordringer. I praksis normalfordeles ikke karakterene, og forventningen om gjennomsnittskarakter medfører lite differensiering i karakterer. I tillegg har bedriften relativt lave finansielle belønninger tilknyttet karakterene, noe som gir lav motivasjonseffekt av karakterene. Bedriftens prestasjonsbaserte aksjeprogram og eksepsjonelle belønninger kompenserer i noen grad for manglende finansielle belønninger. Tabellen under oppsummerer funnene i bedriften.

Oppsummerte funn i bedrift C	
Skjønnsmessige prestasjonsevalueringer	
Helhetlige vurderinger	Ja, tar hensyn til støy og kompleksitet i arbeidsoppgaver.
Observasjon av innsats	God. Alltid to ledere som evaluerer ansatte, og de henter innspill fra flere. Ansatte kan klage om de ikke føler seg godt observert.
Tilbakemeldinger	Flinke til å gi ros, men kvier seg for kritiske tilbakemeldinger.
Trynefaktor	Ansatte og ledere ser at trynefaktor kan påvirke vurderingene, men stort sett ikke et utstrakt problem. Dempes når det er flere som evaluerer prestasjonene og fordi ansatte har mulighet til å klage.
Motvilje mot å differensiere	Ja, ledere finner det krevende å gi svake karakterer. Overvurderer ofte underytere sine prestasjoner.
Karakterer og tvungen fordeling	
Sikre faglig og personlig utvikling	Ja, men er for stort fokus på karakterer heller enn tilbakemeldinger. Karakterer og kalibrering fortrenger jevnlige tilbakemeldinger.
Transparent link mellom ytelse og belønning	Ja, karakterene innebærer omfattende dokumentasjon.
Ytelsesstandarder	Kalibrering sikrer kommunikasjon mellom ledere, det gir likere ytelsesstandarder. Likevel utøves skjønn forskjellig.
Identifisere sterke, svake og normale ytere	Ja, men det oppleves krevende å få svake karakterer.
Differensiering	Nei, det differensieres for lite i karakterer. Karakterene 1 og 5 brukes nesten ikke. Skyldes forventning om gjennomsnittskarakter. Ledere finner det vanskelig å identifisere svake ytere, kan da heller ikke premiere gode prestasjoner. Ansatte finner det demotiverende.
Rettferdig karaktergivning	Delvis. Systemet støtter ikke teamytelse, og enkelte ledere gir karakterer etter hva som passer i fordelingen.
Karakteren 3	Skal være god, men ansatte finner den demotiverende. For mange får karakteren 3.
Målsetting	Insentiv til lite ambisiøs målsetting.
Finansielle belønninger tilknyttet karakterene	Lønnsøkningene tilknyttet karakterene oppleves for små til å motivere. Dette kompenseres delvis for av aksjeprogram og eksepsjonelle belønninger. Har mistet noen av ansatte som får nest høyeste karakter, men stort sett lav grad av turnover.
Nedgangstider	Kuttet prestasjonsbaserte lønnsøkninger.
Ikke-finansielle belønninger	Viktig, men kompenserer ikke for manglende finansielle belønninger.

Tabell 8: Oppsummering av funn i bedrift C

6. Analyse

I dette kapittelet vil de mest sentrale funnene fra de tre casene sammenlignes og drøftes i lys av eksisterende teori og empiri. Hensikten er å skape forståelse for hvordan fastsettelse av karakterer basert på skjønn og tvungen fordeling virker i de ulike bedriftene. Vi bruker, som beskrevet i del 2.4.4, tvungen fordeling som en samlebetegnelse for når bedrifter har fastsatte prosentandeler for hvor mange ansatte kan få hver karakter, og når de har forventning om normalfordeling og gjennomsnittskarakter. Videre omtaler vi ledere med evalueringsansvar gjennomgående som ledere, og ikke-ledende ansatte uten evalueringsansvar som ansatte.

Gjennom analysen vil vi koble momentene opp mot formål med belønningssystemene, bruk av karakterer, kalibrering og tvungen fordeling. Vi anser det derfor hensiktsmessig med en oppsummerende tabell for de ulike formålene. I kapittel 4 og 5 så vi at bedriftene har tilsvarende formål med systemene, og det er derfor egnet med en helhetlig tabell for samtlige bedrifter.

Formål med bedriftenes systemer for belønning og prestasjonsevaluering	
Belønningssystem	Motivere til økt innsats Tiltrekke og beholde ønsket arbeidskraft
Karakterer	Motivere til personlig og faglig utvikling Sikre transparent link mellom belønning og ytelse
Kalibrering	Skape like ytelsesstandarder på tvers av bedriften Sikre åpenhet og kommunikasjon mellom ledere
Tvungen fordeling	Identifisere sterke, normale og svake ytere Sikre differensiering

Tabell 9: Oppsummering av formålene med bedriftenes systemer for belønning og prestasjonsevaluering

I del 6.1. drøftes muligheter og utfordringer med å bruke skjønn i prestasjonsevalueringer. Deretter vil vi i del 6.2. drøfte virkningen av ulike former for tvungen fordeling av ansattes karakterer. Her vil vi også se tvungen fordeling og skjønnsmessige vurderinger under ett. Samlet besvarer del 6.1 og 6.2 utredningens problemstilling:

Hvordan virker prestasjonsevalueringer basert på skjønn og tvungen fordeling av ansattes karakterer?

6.1 Skjønnsmessige prestasjonsevalueringer

Samtlige bedrifter benytter skjønn i prestasjonsevalueringer av ansatte. Vi observerer at ledere og ansatte i de ulike bedriftene forteller om tilsvarende muligheter med skjønnsmessige prestasjonsevalueringer, men skiller seg fra hverandre i hvilke utfordringer de møter og hvor store utfordringene synes å være.

Muligheter med skjønnsmessige prestasjonsevalueringer

Helhetlig vurdering

Ansatte og ledere i alle bedriftene påpeker en rekke fordeler med skjønnsmessige prestasjonsevalueringer. Blant annet gir bruken av skjønn mulighet for mer helhetlige vurderinger, og er avgjørende for å fange opp nyanserikdommen i ansattes arbeidshverdag. Våre observasjoner samstemmer med Gibbs et al. (2004) som finner skjønnsmessige vurderinger egnet, da lederen kan ta hensyn til all relevant informasjon vedrørende ansattes prestasjoner. Bedriftene synes å bruke skjønn fordi ansattes komplekse arbeidsoppgaver ikke kan kvantifiseres i objektive måltall. Skjønnsmessige vurderinger fremstår derfor nødvendig om prestasjonsevalueringene skal reflektere faktisk bidrag til verdiskapningen. Hayes og Schaefer (2000) finner tilsvarende, og beskriver bruk av subjektivitet egnet når deler av prestasjonen ikke er observerbar med kvantitative prestasjonsmål.

I samtlige bedrifter evalueres ansatte både på atferd og leveranse. Ingen av bedriftene nevner problemer med vridningseffekter, til tross for at ansatte har flere og komplekse oppgaver. Det kan tyde på at skjønnsmessige vurderinger av både atferd og leveranse er egnet for å redusere problemer med multitasking og vridningseffekter, slik Bushman, Indjejikian og Smith (1996) hevder. I bedriftene fremstår evaluering av samarbeidsevner særlig viktig. Dette fordi enkeltindivider sjeldent kan gjennomføre sine arbeidsoppgaver på en god måte uten innspill fra kollegaer. Bedriftene ønsker derfor gjennom skjønnsmessige vurderinger å oppmuntre til samarbeid og godt arbeidsmiljø. Ingen av lederne forteller om problemer med samarbeid blant sine ansatte, og ansatte ser på samarbeid som en selvfølge. Skjønnsmessige vurderinger synes på den måten å fungere godt for å sikre at også oppgaver som er vanskelige å kvantifisere blir prioritert (Prendergast, 1999).

Ukontrollerbare forhold

Ansatte i bedriftene opplever skjønnsmessige vurderinger som mindre risikofylte enn objektive måltall. Dette skyldes at deres resultater ofte påvirkes av ukontrollerbare forhold

som ledere kan ta hensyn til når prestasjonsevalueringene baseres på skjønn. Ansattes opplevelser samsvarer med Murphy og Oyer (2003), som hevder skjønnsmessige vurderinger reduserer ansattes risiko fordi støy kan filtreres ut. Gibbs et al. (2004) finner at subjektivitet særlig brukes for å redusere ansattrisiko i nedgangstider og når ansatte har vanskelige mål. Bedriftene vi har undersøkt brukte skjønn i prestasjonsevalueringer også før nedgangstidene, men det har vært spesielt viktig de siste årene for å se den enkeltes innsats i krevende tider. Subjektive prestasjonsmål virker derfor hensiktsmessig for å holde ansatte motiverte når støy påvirker resultatene. Ansatte opplever også at ledere evaluerer hvordan de responderer på støy, da de ofte har delvis mulighet til å dempe negative effekter eller utnytte positive effekter av støyet. Det er derfor rimelig å anta at bedriftene bruker subjektivitet for å redusere risiko ved ukontrollerbare forhold, men samtidig gi insentiv til å respondere så godt som mulig (Gibbs et al., 2004).

I likhet med funnene til Merchant (1989), virker det ikke som bedriftene filtrerer ut støy som påvirker ansattes resultater positivt på samme måte som de filtrerer ut negativt støy. I gode tider fikk ansatte store finansielle belønninger, selv om innsatsen var lavere, fordi positivt støy medføre gode resultater. Blant annet forteller ansatte og ledere i bedrift A om tilfeller ansatte blir belønnet for vellykkede prosjekter selv når de skyldes flaks heller enn ekstraordinær innsats. En mulig forklaring på mindre filtrering av positivt støy, er antakelsen om at risikoaverse ansatte er mer opptatt av å bli forsikret mot nedsiderisiko (Gibbs et al., 2004). Det virker dermed som bedriftene bruker subjektivitet mer for å filtrere ut uflaks enn flaks.

Samlet virker fordelene med skjønnsmessige prestasjonsevalueringer å være muligheten det gir for helhetlige og nyanserte prestasjonsevalueringer, evaluering av ikke-kvantifiserbare forhold som inngår i prestasjonen og reduksjon i risiko når ukontrollerbare forhold påvirker resultatene.

Utfordringer med skjønnsmessige prestasjonsevalueringer

I henhold til Holmstrom (1979) gir effektive prestasjonsmål nøyaktige indikasjoner på individets bidrag til bedriftens verdiskapning og innebærer lav risiko for den ansatte. Utfordringen med skjønnsmessige prestasjonsevalueringer er subjektiviteten i vurderingene, som kan redusere nøyaktigheten av prestasjonsevalueringen og dermed øke risikoen for de ansatte. Flere ansatte i bedriftene forteller om situasjoner de har følt seg urettferdig behandlet når leders skjønnsmessige oppfatning avgjør belønningen de får. Blant annet Prendergast og

Topel (1993) og Fulk, Brief og Barr (1985) har observert tilsvarende i sin forskning. Rettferdighet og tillit i prestasjonsevalueringene synes å ha sammenheng med leders evne til å observere ansattes prestasjonene, gi gode tilbakemeldinger, egenskaper ved lederen og relasjonen mellom leder og ansatt. Slike momenter stemmer overens med de Murphy (2008) finner vedrørende rettferdige og suksessfulle prestasjonsevalueringer. I det følgende vil vi drøfte utfordringer bedriftene møter når prestasjonsevalueringene baseres på skjønn.

Observasjon av innsats

En gjennomgående observasjon i bedriftene er at leders evne til å observere ansattes innsats og prestasjoner er avgjørende for hvordan ansatte opplever skjønnsmessige prestasjonsevalueringer. Dette er i tråd med Murphy (2008), som mener hovedproblemet med subjektive målinger er at ledere verken har motivasjon eller kompetanse til å foreta nøyaktige vurderinger. Problemet er ikke nødvendigvis et ønske om å foreta unøyaktige vurderinger, men informasjonsmangel og kognitive skjevheter som fører til ubevisste og systematiske feilvurderinger (Murphy, 2008).

I bedrift A erfarer ansatte at leder som regel har god innsikt i deres arbeidshverdag. Det oppleves imidlertid utfordrende i tilfeller det ikke er gjeldende. Når leder ikke har tilstrekkelig evalueringsgrunnlag, skal innspill hentes fra andre med bedre mulighet til å observere ansattes prestasjoner. Imidlertid oppfatter ansatte andres innspill som mindre nyanserte, og dermed dårligere egnet til å gi et helhetlig bilde av ansattes prestasjoner. Bruk av andres observasjoner kan derfor kun kompensere til en viss grad om leder har manglende informasjon vedrørende ansattes prestasjoner.

Ansatte i bedrift B har blandede erfaringer med leders mulighet til å observere deres innsats. Som hovedregel skal kun nærmeste leder evaluere ansattes prestasjoner. Hvorvidt ansatte føler seg sett vil derfor avhenge av lederens faglige og fysiske distanse. For ansatte med dualrapportering er den fysiske avstanden til leder med faglig kompetanse gjerne stor, mens lederen som sitter fysisk nært ofte har en annen faglig bakgrunn. Vi opplever derfor at særlig ansatte med dualrapportering finner det utfordrende å bli sett, og i stor grad avhenger av ansattes evne til å fremme egne prestasjoner. I bedrift B kan det virke som Murphy (2008) sine funn er gjeldende for ansatte med dualrapportering, fordi de mener manglende kompetanse kan bidra til mer urettferdige prestasjonsevalueringer.

På en annen side erfarer ansatte uten dualrapportering at nærmeste leder stort sett har god innsikt i deres arbeidshverdag. Det kan blant annet knyttes til at bedriften er av relativt liten størrelse i Norge. Likevel påpeker ansatte at når kun en leder evaluerer, er man avhengig av lederens vilje og mulighet til å skaffe innsikt i deres prestasjoner. Det virker derfor å være forskjeller i lederes evne til å foreta nøyaktige prestasjonsevalueringer. Sammenlignet med bedrift A synes det her å være noe større utfordringer med observasjon av ansattes innsats. Når de som evaluerer ofte har informasjonsmangel og kognitive skjevheter som fører til ubevisste og systematiske feilvurderinger (Murphy, 2008), vil flere ledere i evalueringen kunne gi et mer nyansert og helhetlig bilde av ansattes prestasjoner.

Ansatte i bedrift C uttrykker generelt at ledere har god innsikt i deres prestasjoner. Dette kan ha sammenheng med evaluering av to ledere, samt mulighet til å klage i tilfeller de ikke føler seg riktig oppfattet eller observert. Linjeleder er den som har best innsikt i ansattes prestasjoner, men øvre leder kan tilføye flere nyanser og således gi et bredere vurderingsgrunnlag. Vi opplever klagemuligheten egnet til å bidra til ansattes involvering i prosessen, da de får mulighet til å påvirke hvem som skal vurdere dem ved eventuelle klager. Da også klager fra ansatte med svake prestasjoner tas til etterretning, viser bedriften respekt for samtlige ansatte. Både involvering og respekt er i henhold til Colquitt et al. (2001) avgjørende faktorer for opplevelse av prosedyrerettferdighet. Vi erfarer dermed at klagemuligheten reduserer risikoen for opplevelser av urettferdige og upresise prestasjonsevalueringer i bedrift C.

Ansattes opplevelser i de tre bedriftene tyder på at å hente innspill fra flere i tilfeller nærmeste leder ikke observerer ansatte tilstrekkelig, kun kan kompensere til en viss grad om linjeleder har manglende informasjon om ansattes prestasjoner. Det fremstår derfor avgjørende for ansattes opplevelse av rettferdighet at den som evaluerer faktisk har god innsikt i ansattes arbeidshverdag. Klagemuligheten til ansatte i bedrift C synes i større grad å kompensere for eventuelle svake observasjoner.

Tilbakemeldinger

Locke og Latham (2002) beskriver tilbakemeldinger som avgjørende for ansattes mulighet til å vite hvordan de presterer i forhold til målene, og justere sin atferd deretter. Særlig viktig er dette når målene er kvalitative, fordi slike mål oppleves mindre håndgripelige. Eksempelvis finner ansatte tilbakemeldinger særlig nødvendig om de skal akseptere prestasjonsevalueringen av atferd, da målene på atferd oppleves vagere enn målene på

leveranse. Når skjønnsmessige prestasjonsevalueringer ofte innebærer mindre håndfaste mål enn objektive måltall, er det å gi tilstrekkelig med tilbakemeldinger en sentral utfordring med vurderinger basert på skjønn.

Ansatte i samtlige bedrifter oppfatter i likhet med DeNisi & Kluger (2000) tilbakemeldinger som essensielt for motivasjonen til å prestere godt. De mener også tilbakemeldinger leder til høyere prestasjoner, noe som stemmer med Erez (1977) og Becker (1978) sine funn. I lys av våre observasjoner er leders innsikt i ansattes arbeidshverdag avgjørende om leder skal være i stand til å gi gode og konstruktive tilbakemeldinger. Det er også en sterk sammenheng mellom ansatte som opplever tilstrekkelig med tilbakemeldinger og ansatte som synes leder har god innsikt i deres arbeidshverdag.

I alle bedriftene opplever ansatte stort sett å få jevnlige tilbakemeldinger. På en annen side er det forskjeller mellom ledere, og enkelte ledere kvier seg mer enn andre for å gi tilbakemeldinger. Slike ledere er særlig motvillige mot å gi ærlige tilbakemeldinger når ansatte har prestert under forventning. Dette kan skyldes et ønske om å unngå vanskelige samtaler og eventuelle konflikter. Det er rimelig å anta at ledere her preges av bekvemmelighetshensyn, og unngår å gi tilbakemeldinger som setter dem selv i en vanskelig situasjon. Dette er ikke ubegrunnet, da blant annet Aguinis et al. (2012) hevder negative tilbakemeldinger kan medføre defensive reaksjoner og konflikter. En annen risiko med å gi negative tilbakemeldinger er at de ofte oppleves mindre treffsikre og rettferdige (Aguinis et al., 2012). Vi observerer store forskjeller mellom ledere, der enkelte finner det mer krevende enn andre å gi tilbakemeldinger på prestasjoner under forventning.

Ansatte opplever konkrete og umiddelbare tilbakemeldinger som særlig avgjørende når de presterer under forventet. Dersom ansatte ikke får tilstrekkelig med tilbakemeldinger når prestasjonene ikke tilfredsstiller leders forventninger, gis de heller ikke mulighet til å forbedre sine prestasjoner. Uten denne muligheten opplever ansatte kritiske prestasjonsevalueringer urettferdige. I samtlige bedrifter er målet med jevnlige tilbakemeldinger å unngå misnøye og konflikter i medarbeidersamtalen som følge av overraskende prestasjonsevalueringer. Det er da interessant å merke seg at det er i tilfellene tilbakemeldinger er særlig viktig for ansattes opplevelse av rettferdighet at ledere kvier seg for å gi dem.

Relasjonelle utfordringer

Studier på skjønnsmessige prestasjonsevalueringer viser at egenskaper ved den som evaluerer, og relasjonen mellom den som evaluerer og den som blir evaluert, har like stor betydning for evalueringen som selve prestasjonen til en ansatt (DeNisi & Pritchard, 2006; Levy & Williams, 2004). Dette samsvarer med våre funn, da ansatte i bedriftene vi har undersøkt også opplever at slike forhold har stor påvirkning på prestasjonsevalueringene. En utfordring med skjønnsmessige vurderinger virker dermed å være at prestasjonsevalueringene kan være forårsaket av forhold som ikke burde være relevante for vurderingen, heller enn faktiske prestasjoner

I alle bedriftene opplever ansatte forskjeller i hvordan ledere foretar prestasjonsevalueringer, og hvor stor tillit de har til sine ledere varierer. Grad av tillit synes å avhenge av leders personlig egnethet og erfaring, noe som står i tråd med forskning på området (Gibbs et al., 2004). Ansatte i de tre bedriftene forteller blant annet om store forskjeller i hvordan ledere utøver skjønn, og uttrykker misnøye med at enkelte ledere lettere gir ut gode karakterer enn andre. Vi oppfatter at opplevelsen av rettferdighet utfordres når ledere ikke benytter karakterskalaen på samme måte.

Motvilje mot å gi svake karakterer

Et av formålene med belønningssystemene i alle bedriftene er å motivere ansatte til å yte høy innsats. Ansatte i bedriftene forteller at differensiering i karakterer og belønning er nødvendig om systemet skal virke motiverende. Imidlertid ser vi utfordringer i bedriftene knyttet til differensiering, og ansatte forteller om for lite differensiering i karakterer. Dette synes å være en konsekvens av både kalibreringsprosessen, og at ledere finner det ubehagelig å gi svake karakterer til sine ansatte. Funnene samsvarer med Prendergast (1999), som beskriver lederes motvilje mot å gi svake karakterer som en utfordring med skjønnsmessige vurderinger, da det medfører skjevheter i målingene. Lederes motvilje kan også ha sammenheng med funnene til Bretz og Milkovich (1989), som hevder ledere frykter kritiske prestasjonsevalueringer kan føre til konflikter mellom leder og ansatt.

Ledere i bedrift A og C forteller om en tendens til å unngå å gi svake karakterer. Det kan derfor virke som de preges av ettergivenhets-bias, og overvurderer prestasjonene til svake ytere for å unngå dårlige relasjoner til misfornøyde ansatte (Prendergast, 1999). I tillegg forteller ansatte i bedriftene om lederes tendens til å gi nesten alle ansatte karakterer rundt gjennomsnittet, hvilket innebærer sammenpressing av skalaen (Prendergast, 1999). Dette

tydeliggjøres ved at 95 % av ansatte i bedrift A mottar en kombinasjon av karakterene 3 og 4. Det kan derfor virke som lederne også preges av sentralbias, der ansattes prestasjoner samles på midten av skalaen (Prendergast & Topel, 1993). Sentralbias reduserer verdien av skjønnsmessige vurderinger, fordi forholdet mellom innsats og belønning blir vagere (Prendergast, 1999). Vi oppfatter misnøye blant ansatte knyttet til mangel på differensiering. De finner det urettferdig når ansatte med store forskjeller i prestasjoner ender med samme karakter. Det er rimelig å anta at hvor motiverende karakterene og de skjønnsmessige vurderingene virker, avhenger av hvorvidt de oppleves rettferdige. I så måte kan urettferdighet knyttet til mangel på differensiering innebære at ledere undergraver formålet med belønningssystemet, som nettopp er å motivere de ansatte.

I bedrift B opplever ledere også det ubehagelig å gi svake karakterer, og i tilfeller det er mulig vipper ansatte opp fra karakteren "under forventet". Imidlertid synes problemet mindre enn i bedrift A og C, og ledere virker mer villige til å differensiere i karakterer.

Ingen ansatte eller ledere vi har vært i kontakt med forteller om utfordringer med underrapportering. Vi får opplyst at karakterskalaen brukes på samme måte selv i nedgangstider, og ledere virker ikke å ha insentiv til å underrapportere ansattes prestasjoner. Prendergast (1999) hevder underrapportering sjeldent er et problem i større bedrifter der ledere ikke har krav på bedriftens resultater, noe som samsvarer godt med våre observasjoner.

Trynefaktor

I likhet med Prendergast og Topel (1993) påpeker ansatte i alle bedriftene at skjønnsmessige prestasjonsevalueringer kan være problematiske fordi lederes evaluering kan påvirkes av relasjonen de har til sine ansatte. Det er imidlertid forskjeller i hvor stort problem det virker å være i bedriftene. I bedrift A synes trynefaktor i enkelte tilfeller å få spillerom, men vi oppfatter det ikke som et utstrakt problem i bedriften. Dette kan skyldes at det sjeldent er en enkelt leder som foretar evalueringene.

I bedrift B fremstår trynefaktor som et større problem for de ansatte enn i bedrift A. Det kan skyldes at det kun er en leder som evaluerer deres prestasjoner, og lederens personlig oppfatning får dermed større betydning. Utfordringen med trynefaktor kan tenkes å særlig gjelde for ansatte med dualrapportering. Leder som evaluerer dem har i mindre grad mulighet til å observere deres prestasjoner, og det er da naturlig at trynefaktor spiller en

større rolle i evalueringen. En annen mulig forklaring kan være store forskjeller i finansiell belønning avhengig av prestasjonsevalueringene. Det er ikke nødvendigvis slik at trynefaktor påvirker evalueringene mer i bedrift B enn A, men at ansatte opplever det mer urettferdig fordi det har større konsekvenser. Tillit til leders prestasjonsevaluering virker derfor enda viktigere når evalueringen har store økonomiske konsekvenser for de ansatte.

I bedrift C opplever ikke ansatte trynefaktor som et utstrakt problem. Dette kan skyldes at det alltid er to ledere som foretar vurderingene. På den måten vil aldri en enkelt leder sin personlige oppfatning alene være avgjørende for prestasjonsevalueringen. Videre kan det tenkes at ansatte ikke opplever trynefaktor som et problem fordi de har mulighet til å klage og be om vurdering fra andre om de føler seg urettferdig behandlet. Det virker som ansatte stort sett har tillit til leders prestasjonsevalueringer og opplever dem rettferdige, da klagestatistikken i bedriften er lav. På en annen side kan lav klagestatistikk tyde på at ansatte ikke bruker muligheten i stor grad. Likevel kan man anta at klagemuligheten har en dempende effekt på misnøye tilknyttet trynefaktor, fordi det er ansatte selv som tar stilling til om de godtar prestasjonsevalueringen. Bragelien (2005) støtter opp under resonnementet, og hevder klagemulighet fungerer som en sikkerhetsventil for ansatte i prestasjonsevalueringer.

I alle bedriftene kan ansatte tro det er trynefaktor involvert i prestasjonsevalueringene, selv når det ikke er det. Noen vil alltid oppleve å komme dårlig ut og å ha blitt urettferdig behandlet, til tross for at evalueringen faktisk reflekterer deres prestasjoner. Dette gjelder fordi mennesker har en iboende tendens til å overvurdere egen betydning i bedriften (French, 1968).

Ansatte i de tre bedriftene finner det vanskeligere å bli evaluert på atferd enn leveranse. Årsaken kan være at uklare atferdsmål fører til mer subjektive og uhåndgripelige vurderinger, der trynefaktor oppleves å få mer spillerom. Det ser ut til at ansatte opplever vurderingen av atferd som en indikasjon på hvor godt ledere liker de ansatte, fremfor reelle vurderinger av atferdsmål. Ansatte virker derfor å ha mindre tillit til leders vurdering av atferd, og vurderingen oppleves ofte mindre rettferdig. Dahle (2014) finner tilsvarende, og hevder evaluering av atferd egentlig er en formalisering av trynefaktoren.

I samtlige bedrifter opplever ledere det også mer utfordrende å evaluere og følge opp atferd enn leveranser. Som følge av mer krevende atferdsevalueringer synes lederes evalueringer av leveranser ofte å legge føringer for vurderinger av atferd. Når ledere lar gode eller dårlige

oppfatninger av ansattes leveranser få betydning for evaluering av atferd, preges ledere av halo- og horneffekter (Thorndike, 1920). Halo- og horneffekter i prestasjonsmåling oppstår når en leder registrerer positive trekk ved en ansatt og lar disse påvirke andre deler av vurderingen.

Oppsummert om skjønnsmessige prestasjonsevalueringer

Til tross for utfordringer med skjønnsmessige prestasjonsevalueringer, oppleves bruk av skjønn nødvendig for å se helheten i ansattes prestasjoner. Dette underbygges av at ingen ledere eller ansatte i bedriftene har ytret et ønske om å fjerne skjønn fra prestasjonsevalueringene. Å eliminere relasjonelle utfordringer som trynefaktor og ulik utøvelse av skjønn blant ledere synes umulig. Utfordringene kan imidlertid kompenseres for ved å sikre god observasjon av ansattes innsats, tilstrekkelig med tilbakemeldinger og å gi klagemulighet til de ansatte.

6.2 Tvungen fordeling

Vi vil i det følgende drøfte muligheter og utfordringer med ulike former for tvungen fordeling av ansattes karakterer. Ved bruk av tvungen fordeling ligger det implisitt at bedriftene bruker karakterer. Det er derfor naturlig å også drøfte muligheter og utfordringer med karakterer i prestasjonsevalueringer. I tillegg vil vi se skjønnsmessige prestasjonsevalueringer og tvungen fordeling under ett.

Bedriftene vi har undersøkt benytter alle tvungen fordeling av ansattes karakterer, men systemene er utformet på ulike måter og det er forskjeller i hvor store finansielle belønninger som er tilknyttet karakterene. Bedrift A og C har en forventning om normalfordeling og gjennomsnittskarakter, med relativt lave finansielle belønninger knyttet til karakterene. Bedrift B har derimot fastsatte prosentandeler for hvor mange som kan få hver karakter, med store finansielle belønninger knyttet til karakterene. Våre funn tyder på at hvordan systemene for prestasjonsevaluering virker, varierer ut fra hvordan systemene er utformet og hvor store finansielle belønninger som er knyttet til karaktergivningen.

Ytelsesstandarder og beslutningsgrunnlag

I alle bedriftene oppfattes tvungen fordeling som egnet til å skape likere ytelsesstandarder på tvers av forretningsenheter. Gjennom kalibreringsprosesser tvinges ledere på tvers av avdelinger til å kommunisere rundt karaktersetting. Målet med kalibreringsprosessen er å

fremme en felles forståelse for hvordan karakterskalaen brukes, og sikre rettferdig karaktergivning uavhengig av hvilken leder man har. Også Stewart et al., (2010) finner i sin studie tvungen fordeling og kalibrering egnet til å skape like ytelsesstandarder. På den andre siden ser vi i alle bedriftene forskjeller i hvordan ledere utøver skjønn, der enkelte lettere gir ut gode karakterer enn andre. Når karakterene fastsettes basert på skjønn er det imidlertid urealistisk å anta at alle ledere vil vurdere ansatte på samme måte. Klare retningslinjer og kalibrering kan dermed bidra til å redusere problemer med urettferdighet og ulike ytelsesstandarder, men ikke eliminere de helt.

I tillegg forteller bedriftene at tvungen fordeling av karakterer bidrar til å identifisere sterke, normale og svake ytere. Identifiseringen er viktig i beslutninger vedrørende forfremmelser og utviklingsbehov, noe som er i tråd med det Stewart et al., (2010) finner. På en annen side kan identifiseringen gå på bekostning av ansattes motivasjon, særlig i bedrift A og C. Vi får inntrykk av at ansatte finner det noe demotiverende med karakterer, fordi karakterene oppleves som merkelapper. Våre funn samsvarer med funnene til Kuvaas og Dysvik (2012), som hevder summative evalueringer fungerer til å sortere ansatte i ulike kategorier og til å vurdere tidligere prestasjoner, men ikke til å motivere til fremtidig utvikling og læring. I bedrift A underbygges dette av ledere som forteller om situasjoner der ansatte som har fått svake karakterer har endt opp sykemeldte, eller sluttet, som følge av prestasjonsevalueringene. I bedrift B nevnes ikke utfordringen med merkelapper spesifikt av noen ledere eller ansatte. Dette var overraskende, da blant annet Dahle (2014), i motsetning til oss, finner verbale merkelapper mer krevende enn tallkarakterer.

Transparent link mellom ytelse og belønning

Ledere i de tre bedriftene opplyser alle om hvordan karakter- og fordelingssystemene bidrar til å sikre klare og formelt etablerte ytelseskriterier, med direkte link til ansattes belønning. På den måtes sikres ansattes forståelse for sin belønning og hvordan de presterer i forhold til andre (Stewart et al., 2010). I bedrift A og C forteller blant annet ledere og ansatte om hvordan karakterene innebærer omfattende dokumentasjon av prestasjonsevalueringene som ligger til grunn for belønningen. Dokumentasjonen gjør karakterene, med tilhørende belønninger, enkle å forstå. Våre funn tyder dermed på at karakterene skaper en mer transparent link mellom belønning og ytelse, og gjør evalueringene mer håndgripelige for ansatte.

Selv om karakterer skaper en mer transparent link mellom ytelse og belønning, er jevnlige tilbakemeldinger sentrale om ansatte skal forstå grunnlaget for mottatt belønning. Ledere hevder bruk av karakterer åpner for kommunikasjon, da det sikrer minimum en formell samtale i året. Systemene er utformet for å gi verdifulle og pålitelige tilbakemeldinger på innsatsen til de ansatte. Ved implementering av tvungen fordeling av karakterer, vil det i henhold til Stewart et al. (2010) gjennom kalibrering legges til rette for åpen kommunikasjon vedrørende ansattes prestasjoner og hva de må gjøre for å forbedre prestasjonene. I samtlige bedrifter er målet med jevnlige tilbakemeldinger å unngå misnøye og konflikter i medarbeidersamtalen som følge av overraskende prestasjonsevalueringer.

Imidlertid avdekker vi store forskjeller i hvor flinke ledere er på å gi jevnlige og gode tilbakemeldinger. HR-ledere i alle bedriftene forteller blant annet at stort fokus på karaktergivning går på bekostning av jevnlige og konstruktive tilbakemeldinger. Dette er i motsetning til Blume et al. (2009) og Steweart et al. (2010), som hevder tvungen fordeling stimulerer til åpenhet. Det kan virke som bruk av karakterer og tvungen fordeling retter lederes oppmerksomhet mot evaluering og karaktergivning, fremfor å gi jevnlige tilbakemeldinger i løpet av året. I tillegg oppfatter vi at ledere kvier seg for å gi tilbakemeldinger, da de frykter ansatte oppfatter tilbakemeldingene som en indikasjon på hvilken karakter de vil få. Dette er særlig krevende fordi ledere ikke vet utfallet av kalibreringsprosessen, og derfor ikke ønsker å gi ansatte falske forhåpninger eller unødvendige skuffelser med tilbakemeldingene.

Videre har ansatte også et for stort fokus på karakterer, og mange vektlegger karakteren mer enn tilbakemeldingene som gis. Når formålet med karakterene er å motivere ansatte til faglig og personlig utvikling, er det uheldig at tilbakemeldinger som er sentrale for utviklingen hemmes som følge av kalibreringen.

Differensiering

I likhet med teori vedrørende skjønnsmessige prestasjonsevalueringer (Gibbs et al., 2004; Prendergast, 1999) finner ledere i bedriftene det krevende å differensiere ansattes karakterer. Dette fordi de frykter svake prestasjonsevalueringer kan gå på bekostning av relasjonene de har til sine ansatte. I alle bedriftene tyder våre funn på et ønske om å bruke tvungen fordeling av karakterer for å sikre tydeligere differensiering mellom sterke og svake ytere. Eksisterende teori og empiri finner nettopp mangel på differensiering som bakgrunn for å innføre tvungen fordeling (Blume et al., 2009; Lazear & Gibbs, 2015;

McBriarty, 1988; Stewart et al., 2010). Forskere hevder konsekvensen av tvungen fordeling er større spredning i karakterer. Våre funn tyder derimot på at hvorvidt tvungen fordeling fungerer til å tydeliggjøre forskjellen mellom sterke og svake prestasjoner, avhenger av hvilken type tvungen fordeling som brukes.

For ansatte i de tre bedriftene er differensiering i karakterer avgjørende om de skal virke motiverende. Ved lite differensiering forteller de om mindre insentiver til å yte ekstra innsats. Ledere hevder systemene er teoretisk egnet til å sikre spredning i karakterer, men de blir ikke brukt som tiltenkt. Både ledere og ansatte ytrer et ønske om flere karakterer å skille mellom, da de mener det ville gjort differensiering enklere.

Forventningen om normalfordeling av karakterer i bedrift A og C er ment å sikre at hele karakterskalaen benyttes. Imidlertid ser vi mangel på differensiering av karakterer i de to bedriftene. Eksempelvis i bedrift A mottar 95 % av de ansatte en kombinasjon av karakteren 3 og 4, noe som tyder på små skiller mellom sterke og svake prestasjoner. Det kan skyldes at samtidig som det forventes en normalfordeling, forventes gjennomsnittskarakter på 3. Med forventning om gjennomsnittskarakter kan en leder ikke premiere en ansatt med en god karakter, uten å samtidig gi en annen ansatt en dårlig karakter. Når ledere finner det krevende å identifisere svake ytere, både som følge av kompetansesterke ansatte og som følge av motvilje mot å gi svake karakterer, blir det også vanskelig å premiere gode prestasjoner. Konsekvensen blir mangel på differensiering, der nesten alle ansatte ender med karakterer rundt gjennomsnittet. Det kan for mange fremstå urettferdig når ansatte med svært ulikt prestasjonsnivå ender med samme karakter. Utfordringen virker særlig å gjelde for ledere av mindre grupper, da normalfordeling sjeldent reflekterer individuelle prestasjoner i slike tilfeller (Abelson, 2001). Ledere av mindre grupper differensierer derfor mindre enn ledere av større grupper med ansatte.

Utfordringen med skjønnsmessige vurderinger knyttet til lederes feilvurderinger og motvilje mot å differensiere, synes dermed ikke å dempes av tvungen fordeling i bedrift A og C slik eksisterende litteratur hevder (Blume et al., 2009; Stewart et al., 2010), men heller forsterkes. Det kan skyldes at tidligere forsknings stort sett har undersøkt systemer med fastsatte prosentandeler som skal få hver karakter, heller enn forventet gjennomsnittskarakter og normalfordeling. Våre observasjoner tilsier at forventningen om normalfordeling domineres av forventningen om gjennomsnittskarakter når det gjelder differensiering i karakterer. Det er her viktig å merke seg at bedrift A og C i praksis ikke

følger hovedkontoret sine retningslinjer for karaktergivning, og dermed ikke utøver tvungen fordeling som tiltenkt.

I bedrift B brukes en form for tvungen fordeling som innebærer at en viss prosentandel av ansatte skal motta hver karakter, og kravene er ufravikelige. I motsetning til bedrift A og C, bruker bedrift B fordelingen gitt fra hovedkontoret som tiltenkt. Også i bedrift B ønsker de ansatte enda større grad av differensiering enn slik det er i dag, men samtidig synes ikke problemet like sterkt som i bedrift A og C. I motsetning til de andre bedriftene har ikke ledere her et valg om å avstå fra å differensiere, i slike tilfeller vil kalibreringsprosessen medføre en fordeling av ansattes karakterer uansett. Bedriften klarer dermed å sikre differensiering i karakterer gjennom tvungen fordeling. I likhet med andre forskere (Blume et al., 2009; Stewart et al., 2010) finner vi derfor at når det er fastsatt hvor mange som kan få hver karakter, vil ettergivenhetsbias og sentralbias ved skjønnsmessige prestasjonsevalueringer dempes.

I tillegg kan ledere i bedrift B fraskrive seg deler av ansvaret for ansattes endelige karaktergivning. Dette skyldes at kalibreringsprosessen skjer etter nærmeste leder gir ansatte karakterangivning i medarbeidersamtalen. Ledere kan derfor være mindre ærlige i medarbeidersamtalen og pakke inn vurderingen. Eksempelvis kan de i tilfeller de vurderer en ansatt sine prestasjoner som ”under forventet” velge å i stedet gi uttrykk for at den ansatte vipper mellom “under forventet” og “som forventet” i medarbeidersamtalen, og si at kalibreringsprosessen avgjør den endelige karakteren. På den måten kan ledere fraskrive seg deler av ansvaret for karakterfastsettelsen. Kalibreringsprosessen i etterkant demper derfor frykten for at utgivelse av svakere karakterer skal gå på bekostning av relasjonene. Når ledere har behov for noen andre å dele ansvaret med de når de gir svakere karakterer, virker det likevel som de i utgangspunktet finner det ubehagelig å differensiere. Det er også merkverdig at enkelte ledere i bedrift B bruker systemet med tvungen fordeling som en unnskyldning til å gi ansatte svakere karakterer, heller enn å ta ansvar for å hjelpe svake ytere til å forbedre seg. Blant annet Boyle (2001) finner det uheldig når ledere trenger et system som tvinger frem en differensiering, heller enn å ta ansvar for egen karaktergivning.

Karakterene reflekterer ikke faktiske prestasjoner

Som nevnt ønsker bedrift A normalfordeling i karakterer for å skape større spredning. Imidlertid er lite sannsynlig at individuelle prestasjoner er normalfordelte (Abelson, 2001). Vi oppfatter blant annet at enkelte ledere blir fristet til å gi karakterer til ansatte basert på hva

som passer i kalibreringen, heller enn karakterer som reflekterer faktiske prestasjoner. I så måte blir linken mellom ytelse og belønning mindre transparent, og prestasjonsevalueringen reflekterer ikke lenger faktiske prestasjoner. I henhold til Baker et al. (1994) og Holmstrom (1979), er prestasjonsmål optimale om de reflekterer en medarbeiders faktiske bidrag til verdiskapningen. Om prestasjonsmålet forvrenges for å passe inn i en forhåndsbestemt fordeling vil de derfor ikke lenger være effektive å bruke. I så måte kan forventning om normalfordeling gå på bekostning av rettferdige og nøyaktige prestasjonsevalueringer. Problemet er at fordelingen av karakterer ikke samsvarer med fordelingen av ansattes prestasjoner (Stewart et al., 2010)

I bedrift B fremstår ovennevnte utfordring enda sterkere enn i bedrift A. Guralnik og Wardi (2003) hevder en konsekvens av tvungne prosentandeler kan være at ansatte med like prestasjoner ender med ulike karakterer. Dette skyldes at det i mange tilfeller er flere som burde få samme karakter enn lederne har lov til, og noen må derfor plasseres i andre kategorier enn de egentlig fortjener. Ledere i bedriften uttrykker særlig misnøye med et øvre tak på antall som kan få karakterer over ”som forventet”, da mange ansatte ligger på grensen mellom karakterene ”som forventet” og ”over forventet”. Våre observasjoner tyder på at hvem som vippes opp og hvem som vippes ned i slike tilfeller avgjøres av kalibreringen, men også av tilfeldigheter. Konsekvensen blir at ledere som har brukt opp plassene i de øverste kategoriene tvinges til å gi medarbeidere med tilsvarende prestasjoner ulike karakterer. Dette er en utfordring fordi ansatte som sammenligner seg med andre med tilsvarende prestasjoner, men bedre karakterer, kan oppleve urettferdighet og misnøye (Guralnik & Wardi, 2003).

I likhet med Stewart et al. (2010), finner vi at utfordringen med å bruke en forhåndsbestemt fordeling særlig er krevende for ledere av mindre grupper. Jo mindre gruppene er, desto dårligere egnet er en forhåndsbestemt fordeling av ansattes karakterer (Abelson, 2001). I bedrift B opplever flere ledere sine grupper som særlig talentfulle, der alle ansatte i gruppen presterer i henhold til forventningene eller bedre. Det er derfor vanskelig å identifisere hvem som skal få svakere karakterer. Resultatet kan bli at ansatte i særlig talentfulle grupper straffes om et visst antall av dem må få svakere karakterer enn de ville fått i en annen gruppe (Stewart et al., 2010). I kompetansesterke bedrifter, som de vi har undersøkt, er det rimelig å anta at mange grupper består av dyktige medarbeidere som vanskelig kan forsvare en dårlig karakter. Bedrift B opplyser blant annet at de utelukkende ansetter svært dyktige og erfarne

medarbeidere. Når de da benytter tvungen fordeling av prestasjoner, og en viss andel må få karakteren ”under forventning” selv i tilfeller de ikke har prestert svakt, kan de ansatte oppleve karakterene tilfeldige og gitt for å tilfredsstille fordelingen heller enn å reflektere faktiske prestasjoner (Olson & Davis, 2003). Schrage (2000) underbygger dette, og viser hvordan det blir mindre og mindre nyttig å bruke tvungen fordeling for å sikre differensiering mellom ansatte etter hvert som de blir mer og mer dyktige.

Vi erfarer også at bruk av forhåndsbestemte fordelinger ikke støtter opp under teamytelse når et team som helhet har ytt over gjennomsnittet, da ledere likevel tvinges til å måle de ansatte relativt mot hverandre. Dette er en ulempe med relativ prestasjonsmåling, og medfører at medlemmer av slike team kunne fått bedre karakter om de ble vurdert i en annen gruppe (Olson & Davis, 2003). Tvungen fordeling, særlig slik som bedrift B praktiserer det, skaper således tapere selv i talentfulle grupper (Stewart et al., 2010).

McBriarty (1988) og Truby (2001) hevder bruk av forhåndsbestemte fordelinger kan medføre usunn konkurranse fordi det undergraver. Vi finner imidlertid ikke denne tendensen i noen av bedriftene vi har undersøkt. Verken ledere eller ansatte nevner utfordringer med samarbeid mellom kollegaer, og både ansatte og ledere ser på samarbeid som en selvfølge. Det kan derfor virke som skjønnsmessige evalueringer av både atferd og leveranser fungerer som tiltenkt, ved at også ikke-kvantifiserbare deler av arbeidshverdagen tas hensyn til i vurderingene (Prendergast, 1999). Til tross for at tvungen fordeling ikke støtter opp under teamytelse, bøter bruk av skjønnsmessige vurderinger på denne svakheten med systemene.

Hovedvekten får samme karakter

I bedrift A og C forventes en gjennomsnittskarakter på 3. Karakteren er ment å indikere at man har levert en fullt tilfredsstillende og god jobb. Begrunnelsen er at det i kompetansesterke bedrifter er godt nok å levere som forventet. Ansatte opplever på sin side gjennomsnittskarakteren demotiverende, særlig fordi så mange mottar den selv om prestasjonene er svært ulike. Når et av formålene med belønningssystemet er å motivere ansatte, er det uheldig at karakteren flest ansatte mottar oppleves demotiverende. Ansatte i bedriftene som mottar karakteren 3 mener en høyere karakter bedre ville reflektert deres prestasjoner. Dette samsvarer med både Salamin og Hom (2005) og Zenger (1992), som hevder mennesker har en tendens til å vurdere egne prestasjoner som over gjennomsnittet.

I bedrift B fremstår utfordringen sterkere enn i bedrift A og C. Ledere finner det krevende å skaffe forståelse for at karakteren ”som forventet” skal være god. Det kan skyldes at bedriften kun har fire verbale karakterer. 70-80 % av ansatte skal få karakteren ”som forventet”, og en fast prosentandel kan få hver av de andre karakterene. Karakterskalaen de benytter har færre kategorier enn de andre bedriftene, og ansatte opplever derfor at karakteren ”som forventet” innebærer et for bredt spekter av prestasjoner. Når ”som forventet” omfatter ansatte som har den respektive karakteren på både atferd og leveranse, samt ansatte som har enten “over forventet” eller “under forventet” på et av områdene, blir kategorien for bred til å skille tilstrekkelig mellom gode og dårlige prestasjoner. Både ledere og ansatte har derfor uttrykt et ønske om å dele denne karakteren inn i flere kategorier. Dette gjelder særlig fordi ansatte som har prestert i øvre enden av ”som forventet” opplever karakteren demotiverende.

Utfordringen i bedrift B med å skaffe forståelse fra ansatte om at karakteren ”som forventet” er god, blir forsterket fordi karakterskalaen er asymmetrisk. Det er kun en karakter som er dårligere, men to som er bedre, enn “som forventet”. Karakteren ”under forventet” blir kun gitt til ansatte som virkelig må forbedre både atferd og leveranse, og innebærer tett oppfølging. Det er dermed krevende å få gjennomslag for at karakteren rett over “under forventet” skal oppfattes motiverende av de ansatte.

I likhet med Kuvaas og Dysvik (2012), erfarer vi i alle bedriftene at ansatte blir skuffet når de mottar karakterer som indikerer at de leverer gjennomsnittlig, eller som forventet.

Økte krav over tid

I bedrift A blir utfordringen med at karakteren 3 oppleves demotiverende forsterket når det foreligger krav om økte forventninger til ansatte over tid. Ansatte opplever det stadig vanskeligere å få høyere karakterer enn 3 etter hvert som de blir i bedriften, og at de derfor får mindre og mindre igjen for økt innsats. Når karakteren 3 oppleves lite motiverende i utgangspunktet, er det derfor rimelig å anta redusert motivasjon tilknyttet karakterene jo lenger man jobber i bedriften. En annen utfordring i bedrift A når krav økes over tid, gjelder for ansatte som presterer særlig godt en periode. I slike tilfeller øker ledere sine forventninger til perioden etter, og den ansatte vil motta karakteren 3 ved neste periode om de ikke bedrer sine prestasjoner ytterligere. En mulig effekt kan i følge Brown og Benson (2003) være at ansatte som får toppkarakterer ett år kan bli utmattet. Dette skyldes at forventningene som senere stilles for å få karakterer over 3 vil være så høye at

arbeidsmengden økes utover forsvarlige nivåer. Om funnene til Brown og Benson (2003) stemmer, kan karaktersystemet ha en negativ effekt også på ansatte som får bedre karakterer enn 3.

På den andre siden er det også fordeler med å øke krav til ansattes over tid. For det første kan det bidra til personlig og faglig utvikling av de ansatte, og gi insentiver til å stadig forbedre prestasjoner. For det andre kan justering av krav over tid være hensiktsmessig fordi det gir mulighet til å premiere nyere ansatte som har mer behov for gode karakterer og tilhørende lønnsøkning. De nyere ansatte gjør gjerne samme type jobb som seniorene, men med mindre erfaring og lavere lønnsnivå. Det er derfor naturlig med høyere krav til seniorer enn nyere ansatte. I motsatt tilfelle ville det vært svært krevende for de yngre å få gode karakterer, da ansatte som har jobbet i bedriften over tid har bedre forutsetninger for å lykkes med jobben.

I bedrift B og C er det ikke et system for hvordan ledere skal justere krav over tid. Når ledere her har ulik praksis, kan det skape en opplevelse av urettferdighet for ansatte som ser denne forskjellen mellom ledere. Vi oppfatter at enkelte ledere i de to bedriftene øker krav over tid, og sjeldent gir samme ansatt toppkarakter flere år på rad. Andre ledere gir derimot uttrykk for det motsatte, der typisk de samme får de beste karakterene hvert år. I bedrift B og C kan manglende system for hvordan krav skal økes over tid være et tegn på at kalibrering av ansattes karakterer ikke fungerer til å skape like ytelsesstandarder.

I bedrift B er det særlig utfordrende når enkelte ledere ikke justerer krav over tid, fordi kun en liten prosentandel kan få karakterer høyere enn ”som forventet”. Vi oppfatter at ledere som ikke justerer krav over tid ofte gir de samme ansatte de beste karakterene hvert år. I slike tilfeller synes karaktergivningen i stor grad å baseres på evalueringer fra tidligere år. En konsekvens kan være at ansatte som ikke oppnår karakterer høyere enn ”som forventet” finner systemet demotiverende, da det blir urimelig krevende for dem å forbedre sine karakterer.

Ikke ambisiøs målsetting

En gjennomgående observasjon ved tvungen fordeling av karakterer i bedriftene, er ansattes manglende insentiver til ambisiøs målsetting. Når karakteren man får avhenger av hvorvidt man når målene for året, gis det insentiv til å sette mål som er enklere å nå. Blant annet i bedrift A forteller ledere og ansatte om stadig økende krav for å få gode karakterer, og ansatte som har jobbet i bedriften over tid unngår derfor å ta på seg prosjekter som innebærer

større risiko for å mislykkes, for å sikre måloppnåelse og bedre karakterer. I henhold til Guralnik et al. (2004) medfører slike vridninger mangel på utvikling og forhindret bedriftssuksess. Problemet er at tvungen fordeling gir samtlige ansatte insentiver til å sette enklere mål, også de beste yterne. I så måte er ikke tvungen fordeling av karakterer kun demotiverende for gjennomsnittlige ansatte, men det kan også virke demotiverende for de flinkeste ansatte (Schrage, 2000)

Finansielle belønninger

Prestasjonsevalueringene i bedrift A er kun tilknyttet lønnsøkninger, med unntak av ansatte på ledernivå som også har tilgang til individuell bonus. Generelt virker variasjonen i lønnsøkninger mellom sterkeste og svakeste ytere for små til å ha signifikant betydning på personlig økonomi. Våre funn tyder på at utfordringen med for liten grad av differensiering forsterkes når de få som faktisk skiller seg ut i positiv eller negativ retning ikke får finansielle belønninger som tydeliggjør forskjellen. Med prestasjonsbasert belønning ønsker bedriften at de som faktisk yter mest opplever å få ekstra belønning, for å rettferdiggjøre forskjellen i arbeidsinnsats. Vi oppfatter ansatte som yter mer enn gjennomsnittlig og får toppkarakter misfornøyde når de får belønninger av tilsvarende størrelse som underytere. De poengterer at det er urettferdig med like belønninger for betydelige forskjeller i arbeidsinnsats. Det kan i bedrift A derfor virke som karakterene med tilhørende belønninger er demotiverende for de flinkeste ansatte, da det er rimelig å anta nær tilknytning mellom motivasjon og opplevelse av rettferdighet (Kuvaas & Dysvik, 2012). Når systemet slik det er nå ikke virker motiverende for de som yter mest, er det naturlig å stille seg spørsmål om den egentlige nytten ved å bruke det.

I bedrift C oppfatter vi tilsvarende utfordringer som i bedrift A, da karakterene også her er knyttet til relativt små forskjeller i individuell lønnsøkning. Likevel fremstår utfordringen mindre i bedrift C. Dette kan skyldes ansattes mulighet for prestasjonsbaserte aksjer. I tillegg kan muligheten til å få spontane eksepsjonelle belønninger ved ekstraordinær innsats gjøre de flinkeste ansatte i bedrift C mer tilfredse enn ansatte i bedrift A. Spontane belønninger kan i henhold til Deci et al. (1999) også være mer hensiktsmessige enn forventede belønninger, da de ikke påvirker ansattes indre motivasjon negativt.

Grunnet nedgangstider, og derav mindre mulighet for å belønne ansatte de to siste årene, har bedrift A midlertidig fjernet prestasjonsbaserte lønnsøkninger tilknyttet karakterene. Vi får inntrykk av at både ansatte og ledere opplever bortgangen av prestasjonsbaserte lønnstillegg

frustrerende. Særlig gjelder dette ansatte som har fått gode karakterer. Følgelig får karakterene mindre betydning enn i normalår. Samtidig forstår ansatte bakgrunnen for at prestasjonstillegget blir fjernet når bedriftens rammer er mindre. I tillegg synes ansatte forskjellene i lønnstillegg i normalår er så små at overgangen til nedgangstider ikke har hatt så stor effekt på motivasjonen. Bedrift C har også valgt å kutte prestasjonsbaserte lønnstillegg som følge av nedgangstidene, og har møtt tilsvarende utfordringer som bedrift A.

I bedrift B finner vi ikke samme utfordringer som i bedrift A og C knyttet til finansiell belønning. Dette skyldes store forskjeller i utformingen av belønningssystemet, og at ansatte oppfatter de prestasjonsbaserte belønningene som generøse. En utfordring for dem er at de store forskjellene i finansielle belønninger avhengig av hvilken karakter man får kan oppleves urettferdige. Når kun et fåtall kan få de beste karakterene, vil hvem som ender med disse karakterene i enkelte tilfeller avgjøres av tilfeldigheter. Konsekvensen kan bli at et system som i utgangspunktet er generøst oppleves urettferdig når forskjellene blir så store mellom ansatte med tilsvarende prestasjoner.

På en annen side opplever ansatte i bedrift B store variasjoner i prestasjonsbaserte belønninger som en driver for motivasjon, da de som legger ned ekstra mye innsats faktisk får uttelling for det. Ansatte mener også fastlønnen er konkurransedyktig, slik at uansett hvilken bonus man mottar vil man føle seg godt ivaretatt økonomisk. Bedrift B har i likhet med bedrift A og C valgt å kutte prestasjonsbaserte lønnsøkninger som følge av nedgangstidene, men de har beholdt de individuelle bonusene. Ansatte virker positivt overrasket over bonusene de har mottatt i nedgangstidene, særlig når bonusens størrelse har vært av tilsvarende størrelse som i normalår.

Turnover

I bedrift A forteller ledere at ansatte som typisk får de beste karakterene slutter i bedriften. Det virker dermed som formålet med belønningssystemet om å tiltrekke og beholde talenter ikke ivaretas. Våre observasjoner indikerer at de slutter fordi andre bedrifter i bransjen tilbyr høyere belønning for fremragende prestasjoner. Utfordringen kan ha sammenheng med karaktergivning, da ansatte med fremragende prestasjoner som får karakteren 5 blir bevisste sin verdi i bedriften. Tvungen fordeling er ment å sikre identifisering av sterke, svake og normale ytere, og gi mulighet til å belønne i henhold til prestasjoner (Guralnik & Wardi, 2003). Mangel på tilstrekkelig finansiell belønning tilknyttet karakteren da kan medføre at de

som er identifisert som sterke ytere velger å slutte. Bruk av tvungen fordeling kan dermed kun bidra til å senke turnover av nøkkelpersoner og kritiske talenter om de blir belønnet tilstrekkelig (MacDougall, 1991).

I bedrift C er det lav grad av turnover både i gode tider og nedgangstider. Imidlertid får vi inntrykk av at også denne bedriften har mistet flinke ansatte til andre bedrifter i bransjen som tilbyr høyere lønn. Utfordringen virker mindre her enn i bedrift A, da de opplyser at de ikke mister sine beste ansatte. Dette kan skyldes at slike blir godt ivaretatt med eksepsjonelle belønninger og aksjeprogram. De bedrift C typisk har mistet til andre operatørselskaper har vært de som yter over det som forventes, men likevel ikke får toppkarakter. Slike ansatte virker dermed å få for liten økonomisk kompensasjon for å yte mer enn gjennomsnittet.

I bedrift B virker de finansielle belønningene tilknyttet karakterene positivt på formålet om å tiltrekke og beholde ønsket arbeidskraft. Ledere og ansatte uttrykker at mange har ønsket å starte i bedriften nettopp grunnet den generøse bonusordningen. Funnene stemmer overens med Lazear (2000), som hevder bedrifter med sterke prestasjonsbaserte insentiver vil tiltrekke seg den mest produktive arbeidskraften. I tillegg kan det virke som bonusene brukes som et virkemiddel til å få ansatte til å bli værende i bedriften, noe som samsvarer med det Blakemore et al. (1987) finner i sin forskning. Dette underbygges av at deler av bonusen til nøkkelpersoner først utbetales to år etter den er opptjent.

Samlet finner vi at virkningene av bedriftenes systemer for tvungen fordeling og karakterer avhenger av hvorvidt tilstrekkelige finansielle belønninger er tilknyttet karakterene. Formålet med bedriftenes belønningssystemer er å motivere ansatte til økt innsats, samt tiltrekke og beholde ønsket arbeidskraft. Våre observasjoner tilsier at bedrift B gjennom sine finansielle insentiver er i best stand til å ivareta slike formål. I bedrift A og C fremstår forskjellene i belønning tilknyttet karakterene for små til å ha særlig effekt på ansattes motivasjon. Begge bedriftene har i tillegg mistet viktige ansatte til andre operatørselskaper som i større grad premierer høy innsats.

Både bedrift A og C har prestasjonsbaserte belønninger som er relativt lave, samtidig som det er små forskjeller mellom de som yter mest og de som yter minst. Det er da rimelig å anta at ansatte anser belønningen som et for svakt insentiv. Ved å benytte seg av svake prestasjonsbaserte insentiver i bedrift A og C er det fare for at ansattes indre motivasjon fortrenges, uten at ytre motivasjon øker betydningsfullt (Gneezy & Rustichini, 2000). I så

måte blir totalmotivasjonen lavere. I bedrift B opplever de ansatte derimot finansielle belønninger tilknyttet prestasjoner som betydningsfulle, og virker således mer motiverte både til å yte høy innsats og til å bli værende i bedriften. Våre observasjoner stemmer derfor godt overens med Gneezy og Rustichini (2000) sin konklusjon: ”Pay enough or don’t pay at all”.

Ikke- finansielle belønninger og indre motivasjon

I alle de tre bedriftene opplever ansatte ikke-finansielle belønninger som spennende arbeidsoppgaver, utviklingsmuligheter og anerkjennelse som viktige tillegg til finansiell belønning. Ansattes opplevelser kan ses i sammenheng med Latham og Locke (1979), som mener økonomiske insentiver gjerne er det primære insentivet for ansatte i bedrifter, men uten andre former for belønning ikke er tilstrekkelig for å motivere til høy innsats. Prestasjonsbaserte lønnsøkninger og bonuser burde i henhold til Armstrong (2012) derfor støttes av ikke-finansielle belønninger.

Til tross for at ansatte i bedrift A og C opplever mangel på tilstrekkelige finansielle belønninger, uttrykker ansatte at de opplever arbeidsoppgavene som motiverende i seg selv, og at det er viktig for dem å utvikle sine evner og ferdigheter. Dette kan i henhold til Ryan og Deci (2000) tyde på at ansatte drives av indre motivasjon. Selv om ansatte mener ikke-økonomisk belønning og indre motivasjon i noen grad kompenserer for begrensede finansielle belønninger, fremstår det likevel på sikt nødvendig med økonomisk kompensasjon for gode prestasjoner.

I bedrift B får vi derimot inntrykk av at ansatte mener både finansielle og ikke-finansielle belønninger er godt ivaretatt. De opplever bonusene som veiledende og motiverende, og mener motivasjonen forsterkes av ikke-finansielle belønninger. Det virker dermed som at de ytre belønningene ansatte mottar oppfattes informerende om at de er kompetente, og at denne belønningen ikke går på bekostning av ansattes indre motivasjon. Våre funn samsvarer dermed godt med Deci et al. (1999) sine observasjoner.

Oppsummert om tvungen fordeling av karakterer

Tvungen fordeling av karakterer innebærer både muligheter og utfordringer. I alle bedriftene gir karakterene en mer transparent link mellom ytelse og belønning, samt sikrer identifisering av sterke, normale og svake ytere. På en annen side fortrenger kalibreringen jevnlige tilbakemeldinger som er sentrale for ansattes faglige og personlige utvikling.

I bedriftene vi har undersøkt sikrer bruk av fastsatte prosentandeler for hvor mange som kan få hver karakter differensiering i karaktergivningen, mens forventning om gjennomsnittskarakter forsterker utfordringer knyttet til lav grad av differensiering. En utfordring med alle bedriftenes systemer er at hovedvekten av ansatte mottar karakterer som indikerer prestasjoner i henhold til leders forventninger, men som ansatte ikke motiveres av. Systemene motiverer dermed ikke den største andelen av de ansatte.

En ytterliggere utfordring med tvungen karakterfordeling er at karakterfordelingen ikke nødvendigvis reflekterer faktisk spredning i ansattes prestasjoner. Således kan systemene medføre urettferdig allokering av belønninger. Samtidig er det enklere for ansatte å få gode karakterer om de setter mindre ambisiøse mål, slik at systemene gir insentiver til å unngå prosjekter og oppgaver som er mer krevende. Avslutningsvis finner vi at svake finansielle insentiver knyttet til karakterene kan føre til at bedriftene mister sine flinkeste ansatte. Ansatte som får toppkarakterer blir bevisst sin verdi i bedriften, og uten tilstrekkelig finansielle belønninger er det risiko for at de går de over til bedrifter med bedre mulighet til å gi ekstra kompensasjon for gode prestasjoner.

7. Konklusjon

Vi oppsummerer i det følgende utredningens hovedfunn, begrensninger og forslag til videre forskning. Formålet med denne utredningen er å øke innsikten i hvordan prestasjonsevalueringer basert på skjønn og tvungen karakterfordeling fungerer. Studiens empiriske grunnlag er tre bedrifter i oljebransjen, som alle bruker slike systemer for å evaluere sine ansattes prestasjoner. To av bedriftene benytter en karakterskala fra 1 til 5, med forventning om normalfordeling og gjennomsnittskarakter på 3. Den siste bedriften anvender fire verbale karakterer og forhåndsbestemte prosentandeler for hvor mange ansatte som kan få hver karakter. Grunnet et ønske om dybdeinnsikt på områder med få empiriske studier, har utredningen et eksplorerende forskningsdesign med semistrukturerte intervjuer av både ledere og ikke-ledende ansatte.

Med dette utgangspunkt søker vi å svare på vår problemstilling:

"Hvordan virker prestasjonsevalueringer basert på skjønn og tvungen fordeling av ansattes karakterer?"

7.1 Svar på problemstillingen

Analysen avdekker både muligheter og utfordringer med prestasjonsevalueringer basert på skjønn og tvungen fordeling av ansattes karakterer. Våre funn tyder på at virkningen av slike prestasjonsevalueringer avhenger av systemenes utforming og hvor store finansielle belønninger som er tilknyttet karakterene. I det følgende presenteres de viktigste momentene fra analysen.

Skjønnsmessige prestasjonsevalueringer medfører utfordringer knyttet til subjektiviteten som ligger i leders vurdering av sine ansatte, da lederens personlige oppfatning avgjør hvilken karakter og belønning de ansatte mottar. Evalueringer basert på skjønn kan dermed føre til opplevelse av urettferdighet avhengig av hvilken leder man har. Vi finner blant annet at leders vurdering kan påvirkes av relasjonen de har til sine ansatte, hvor flere ansatte opplever at evalueringene preges av trynefaktor. Ansatte opplever trynefaktor som et større problem når det kun er en leder som evaluerer deres prestasjoner, og når de ikke har mulighet til å klage dersom de opplever evalueringen urettferdig. I tillegg synes ansatte trynefaktor er mer utfordrende når evalueringen har store økonomiske konsekvenser. For at ansatte skal

akseptere og være tilfredse med evalueringen de mottar, finner vi derfor at tillit til leder og opplevelse av rettferdighet i evalueringsprosessen er avgjørende.

Til tross for enkelte utfordringer fremstår skjønnsmessige vurderinger nødvendig for å se helheten i ansattes prestasjoner. Dette underbygges av at ingen ledere eller ansatte i bedriftene ønsker å fjerne skjønn i evalueringsprosessene. Årsaken er at de ansatte har komplekse arbeidsoppgaver som ikke kan kvantifiseres og vurderes med objektive måltall. I tillegg tilrettelegger skjønnsmessige vurderinger for mer nyanserte prestasjonsevalueringer, der støy som påvirker ansattes prestasjoner også kan tas hensyn til. Om ledere har god mulighet til å observere ansattes innsats og gir tilstrekkelig med tilbakemeldinger, samtidig som ansatte har klagemulighet om de føler seg urettferdig behandlet, virker skjønnsmessige vurderinger godt egnet for å evaluere ansattes prestasjoner.

Vår analyse indikerer at tvungen fordeling fortrenger jevnlige tilbakemeldinger som er nødvendige for ansattes opplevelse av skjønnsmessige prestasjonsevalueringer. Enkelte ledere i bedriftene vegrer seg for å gi tilbakemeldinger underveis i året, i frykt for at tilbakemeldingene kan oppfattes som en karakterindikasjon. Samtidig går ledernes fokus på karakterer på bekostning av tilbakemeldingene de samtidig skal gi i medarbeidersamtalen. Dette er uheldig når tilbakemeldinger er avgjørende for ansattes personlige og faglige utvikling.

Ledere uttrykker vanskeligheter med å gi svake karakterer til sine ansatte, fordi det kan medfører konflikter og misnøye. De overvurderer derfor ofte underytere sine prestasjoner i evalueringene. Tvungen fordeling av karakterer er ment å motvirke lav grad av differensiering som følge av motvilje mot å gi svake karakterer. Imidlertid finner vi at tvungen fordeling kun øker grad av differensiering dersom det er ufravikelige krav til gitte prosentandeler av ansatte som kan få hver karakter. Når bedrifter bruker forventning om normalfordeling og gjennomsnittskarakter, åpnes det for at ledere kan avstå fra å differensiere. Dersom ledere ønsker å premiere en ansatt med en god karakter, må en annen ansatt få en tilsvarende svakere karakter om gjennomsnittet skal overholdes. Med lederes motvilje mot å gi svake karakterer blir konsekvensen mangel på differensiering, der nesten alle får karakterer rundt gjennomsnittet. Utfordringen med lav grad av differensiering forsterkes dermed om kravet om normalfordeling ikke er ufravikelig, og det samtidig forventes en gjennomsnittskarakter. Ansatte opplever på sin side mangel på differensiering

urettferdig og demotiverende, fordi flinke ansatte ender med samme karakter og belønning som ansatte med svakere prestasjoner.

Til tross for at tvungen fordeling kan øke grad av differensiering, kan forventningen om gitte fordelinger gå på bekostning av rettferdig karaktergivning. Vår studie indikerer at enkelte ledere gir karakterer basert på hva som passer i fordelingen, heller enn karakterer som reflekterer faktiske prestasjoner. Blant annet forteller ledere i bedriftene at de opplever ansatte som kompetansesterke og produktive, og utelukkende rekrutterer ansatte med høye kvalifikasjoner og kompetanse. Med tvungen fordeling forventes en viss andel likevel å få svake karakterer. Dette kan medføre tilfeldig karaktergivning som ikke nødvendigvis reflekterer ansattes prestasjoner.

Tvungen fordeling støtter heller ikke opp under teamytelse når et team som helhet har ytt mer eller mindre enn det som forventes. Ansatte måles relativt mot hverandre, og karakterene i gruppen forventes å tilfredsstille en forhåndsbestemt fordeling. Ansatte i særlig talentfulle grupper straffes derfor om en viss andel må få svakere karakter enn de ville fått i en annen gruppe. Tvungen fordeling, særlig når tvungne prosentandeler benyttes, skaper derfor tapere selv i talentfulle grupper. Til tross for at tvungen fordeling ikke støtter opp under teamytelse, finner vi ikke tendenser til usunn konkurranse. Dette tilskrives skjønnsmessige vurderinger av både atferd og leveranse, som sikrer at ansattes samarbeidsvilje ivaretas.

Tvungen fordeling kan i tillegg føre til at ansatte med tilsvarende prestasjoner ender med ulike karakterer. Dette gjelder særlig når det benyttes fastsatte prosentandeler. Et øvre tak på antall ansatte som kan få hver karakter, medfører at ledere som har brukt opp sine plasser i enkelte kategorier tvinges til å gi medarbeidere med tilsvarende prestasjoner ulike karakterer. Ansatte som ender med svakere karakterer enn fortjent finner karakterene demotiverende, da de opplever at karakteren ikke reflekterer deres faktiske prestasjoner.

Med forventning om normalfordeling og gjennomsnittskarakter mottar hovedvekten av ansatte gjennomsnittskarakteren. Til tross for at denne karakteren indikerer at man har prestert i henhold til forventningene, opplever ansatte den som lite motiverende. Dette fordi den innebærer et for bredt spekter av prestasjoner. Med tvungne prosentandeler sikres større grad av differensiering. Likevel opplever ansatte også her at karakteren de fleste ansatte skal motta er demotiverende. Problemet blir forsterket når karakterskalaen som brukes ikke er

symmetrisk, da det er færre karakterer som er dårligere enn bedre enn karakteren hovedvekten av ansatte får. I alle bedriftene vi har undersøkt finner vi derfor at hovedvekten av ansatte, enten det benyttes fastsatte prosentandeler eller forventning om gjennomsnittskarakter og normalfordeling, får en karakter de ikke finner motiverende. Tvungen fordeling kan således virke uegnet til å bedre ansattes prestasjoner og arbeidsinnsats.

Analysen indikerer også at størrelsen på finansielle belønninger tilknyttet karakterene er avgjørende for virkningene av prestasjonsevalueringer basert på skjønn og tvungen fordeling. Uten tilstrekkelige finansielle belønninger risikerer man både at karakterene er demotiverende heller enn motiverende, og å miste sine flinkeste ansatte. Sistnevnte skyldes at bruk av karakterer i evalueringene gjør de flinkeste ansatte bevisste sin verdi, hvilket medfører at de krever finansiell kompensasjon for å bli værende i bedriften.

Gneezy og Rustichini (2000) finner i sin studie svekket motivasjon og arbeidsinnsats ved svake finansielle insentiver for oppgaver som preges av høy indre motivasjon. Studien viser på en annen side at innsatsen øker når styrken i de økonomiske insentivene øker. De konkluderer derfor med: "Pay enough or don't pay at all". Deres konklusjon er også gjeldende for vår studie, da vi finner svekket motivasjon som følge av svake finansielle insentiver. Samtidig kan bruk av karakterer med svake finansielle insentiver medføre at bedriftene mister sine flinkeste ansatte. Bruk av tid og ressurser på karaktergivning har dermed ingen hensikt med mindre det differensieres i karakterer og ansatte belønnes deretter. Vår konklusjon er derfor "Pay enough or don't rate at all".

Implikasjoner

Vår studie indikerer at prestasjonsevalueringer basert på skjønn og tvungen karakterfordeling kan fungere for å motivere ansatte, og bidra til å tiltrekke og beholde ønsket arbeidskraft. På en annen side kan bedrifter med slike systemer risikere at hovedvekten av de ansatte blir demotiverte av en gjennomsnittlig karakter, at de beste ansatte slutter uten tilstrekkelig finansiell belønning og at ansatte som får svake karakterer sykemeldes. Hvordan systemene virker avhenger av hvordan de er utformet og hvor store finansielle belønninger som er tilknyttet prestasjonsevalueringene.

Skjønnsmessige prestasjonsevalueringer fremstår nødvendig for å evaluere ansatte med komplekse arbeidsoppgaver. Våre funn indikerer at ansattes tilfredshet med evalueringene

øker om de gis tilstrekkelig med tilbakemeldinger, om lederen har god mulighet til å observere deres prestasjoner, om det er flere ledere som er med i evalueringen av den enkelte ansatte og dersom de ansatte har mulighet til å klage på evalueringen. Hvis bedrifter ønsker å bruke karakterer og tvungen fordeling i prestasjonsevalueringer, tyder våre funn på at det er nødvendig å differensiere mellom sterke og svake prestasjoner, både i belønning og karakterer. I motsatt tilfelle risikerer bedrifter både å miste sine flinkeste ansatte, samtidig som de resterende ansatte blir demotiverte av å ikke få gode karakterer. Vår studie indikerer også at det må foreligge ufravikelige krav til karakterfordeling om den faktisk skal overholdes når ledere gir karakterer til sine ansatte.

Uansett hvilken type tvungen fordeling som brukes er det risiko for karaktergivning basert på hvilke karakterer som passer i fordelingen, heller enn hvilke karakterer som reflekterer faktisk spredning i prestasjoner. Før man innfører slike systemer kan det derfor være hensiktsmessig å undersøke om ansattes prestasjoner faktisk skiller seg fra hverandre på en måte som gjør tvungen fordeling egnet, og om det er behov for å identifisere sterke, normale og svake ytere. Uten behov for dette synes tvungen karakterfordeling å være en risikofylt metode for å evaluere ansattes prestasjoner. Prestasjonsevalueringer basert på skjønn og tvungen karakterfordeling krever mye tid og ressurser, som vanskelig kan rettferdiggjøres med mindre de bidrar til å motivere ansatte og tiltrekke og beholde ønsket arbeidskraft.

7.2. Begrensninger ved studien og forslag til videre forskning

I vår studie har vi avdekket funn som kan undersøkes videre.

Hensikten med utredningen var å skaffe økt innsikt i prestasjonsevalueringer basert på skjønn og tvungen karakterfordeling. Våre funn baseres på 15 intervjuer i tre bedrifter i oljebransjen, i tillegg til tre bakgrunnsintervjuer. Siden vi har gjennomført en eksplorerende casestudie med relativt få intervjuobjekter, foreslår vi at andre gjennomfører lignende studier av bedrifter i samme bransje. Det kan da undersøkes om de finner tilsvarende funn.

Bedriftene vi har studert bruker ulike former for tvungen fordeling. To av bedriftene har en forventning om normalfordeling og gjennomsnittskarakter, mens en bedrift benytter forhåndsbestemte prosentandeler for hvor mange ansatte som kan få hver karakter. De ulike formene for tvungen fordeling synes å innebære ulike muligheter og utfordringer for

bedriftene. Tidligere forskning har hovedsakelig undersøkt tvungen fordeling i form av forhåndsbestemte prosentandeler. Vi anbefaler derfor å gjennomføre flere studier i bedrifter med forventning om gjennomsnittskarakter, for å se om lignende funn avdekkes.

I oljebransjen bruker en rekke bedrifter karakterer og tvungen fordeling i sine prestasjonsevalueringer. Bruk av karakterer og tvungen fordeling er også utbredt i andre bransjer enn den vi har studert. Vi anbefaler derfor å utføre mer forskning i andre kompetansesterke bransjer, for å undersøke om slike bedrifter møter lignende muligheter og utfordringer ved bruk av karakterer og tvungen fordeling.

8. Litteraturliste

Aadland, C. (2015, 25. mars). *Over halvparten av oljebedriftene venter nedgang*. Hentet fra http://www.sysla.no/2015/03/25/oljeenergi/over-halvparten-av-oljebedriftene-venter-nedgang_44995/

Aase, T. H., & Fossåskaret, E. (2007). *Skapte virkeligheter: kvalitativt orientert metode.* Oslo: Universitetsforlaget.

Abelson, R. (2001, 19. mars). *Companies turn to grades and employees go to court*. Hentet fra http://www.nytimes.com/2001/03/19/business/companies-turn-to-grades-and-employees-go-to-court.html

Adams, J. S. (1965). Inequity in Social Exchange. *Advances in Experimental Social Psychology, 2*, 267-299.

Aguinis, H. (2009). *Performance Management* (2. utg.). Upper Saddle River, NJ: Pearson Prentice-Hall.

Aguinis, H., Gottfredson, R. K., & Joo, H. (2012). Delivering effective performance feedback: The strengths-based approach. *Business Horizons, 55*(2), 105-111.

Alchian, A. A., & Demsetz, H. (1972). Production, Information Costs, and Economic Organization. *The American Economic Review, 62*(5), 777-795.

Allen, F., & Gale, D. (1992). Measurement Distortion and Missing Contingencies in Optimal Contracts. *Economic Theory, 2*(1), 1-26.

Amalfe, C. A., & Adelman, H. (2002). *Forced rankings: The latest target of plaintiff 's employment lawyers.* Hentet fra http://www.gibbonslaw.com/news_ publications/articles.php?action=display_publicat ion&publication_id=790

Armstrong, M. (2007). *A Handbook of Employee Reward Management and Practice* (2. utg.). London: Kogan Page.

Armstrong, M. (2012). *Armstrong's Handbook of Reward Management Practice: Improving Performance Through Reward* (4. utg.). London: Kogan Page.

Arrow, K. (1985). The economics of agency. I J. W. Pratt, & R. J. Zeckhauser, *Principals and Agents: The Structure of Business.* Harvard Business School Press.

Baiman, S., & Rajan, M. V. (1995). The informational advantages of discretionary bonus schemes. *The Accounting Review, 70*(4), 557-579.

Baker, G. (1992). Incentive contracts and performance measurement. *Journal of Political Economy, 100*(3), 598-614.

Baker, G. (2000). The use of performance measures in incentive contracting. *The American Economic Review, 90*(2), 415-420.

Baker, G., Gibbons, R., & Murphy, K. J. (1994). Subjective Performance Measures in Optimal Incentive Contracts. *Quarterly Journal of Economics, 109*(4), 1125-1156.

Barney, J. B., & Wright, P. M. (1997). On becoming a strategic partner: The role of human resources in gaining competitive advantage. *CAHRS Working Paper*.

Baron, J., & Kreps, D. M. (1999). *Strategic human resources: Frameworks for general managers.* New York: John Wiley & Sons Inc.

Barth, E. (2005). Prestasjonslønn – kan det bare gå oppover? *Horisont, 6*(4), 114-122.

Bates, S. (2003). Forced ranking. *HRMagazine, 48*(6), 62–68.

Becker, L. J. (1978). Joint effect of feedback and goal setting on performance: a field study of residential energy conservation. *Journal of Applied Psychology, 63*(4), 428-433.

Blakemore, A. E., Low, S. A., & Ormiston, M. B. (1987). Employment Bonuses and Labor Turnover. *Journal of Labor Economics, 5*(4), 124-135.

Blasi, J. R., Freeman, R. B., Mackin, C., & Kruse, D. L. (2008). Creating a Bigger Pie? The Effects of Employee Ownership, Profit Sharing, and Stock Options on Workplace Performance. *NBER Working Paper Series*, 14230.

Blume, B. D., Baldwin, T. T., & Rubin, R. S. (2009). Reactions to Different Types of Forced Distribution Performance Evaluation Systems. *Journal of Business and Psychology, 24*(1), 77-91.

Bol, J. C. (2008). Subjectivity in Compensation Contracting. *Journal of Accounting Literature, 27*, 1-24.

Bommer, W. H., Johnson, J. L., Rich, G. A., Podsakoff, P. M., & MacKenzie, S. B. (1995). On the interchangeability of objective and subjective measures of employee performance: A meta-analysis. *Personnel Psychology, 48*(3), 587-605.

Bose, A., Pal, D., & Sappington, D. E. M. (2010). Equal Pay for Unequal Work: Limiting Sabotage in Teams. *Journal of Economics and Management Strategy, 19*(1), 25-53.

Boyle, M. (2001). Performance reviews: Perilous curves ahead. *Fortune,* 187-188.

Bragelien, I. (2003). Bruk av lønn som styringsinstrument: Hvorfor så mange mislykkes. *Magma - Tidsskrift for økonomi og ledelse, 6*(2), 47-57.

Bragelien, I. (2005). 10 bonustabber – Hvordan lære av teori og praksis? *Praktisk økonomi & finans, 22*(2), 25-35.

Bragelien, I. (2012). Individuelle eller kollektive belønninger? – Teori, empiri og norske bankeksempler. *Praktisk økonomi & finans, 29*(4), 39-52.

Bragelien, I., & Kvaløy, O. (2014). Lønn og belønning. I A. Mikkelsen, & T. Laudal, *Strategisk HRM 2: HMS, Etikk og Internasjonale perspektiver* (s. 328-355). Oslo: Cappelen Damm Akademisk.

Bretz, R. D., & Milkovich, G. T. (1989). *Performance appraisal in large organizations: Practice and research implications.* New York: Cornell University.

Brewster, C., & Larsen, H. H. (2000). *Human Resource Management in Northern Europe: Trends, Dilemmas, and Strategy.* Oxford: Blackwell Publishing.

Brown, M., & Benson, J. (2003). Rated to exhaustion? Reactions to performance appraisal processes. *Industrial Relations Journal, 34*(1), 67-81.

Buchanan, D. A. (2010). *Organizational behaviour* (7. utg.). Harlow: Financial Times/Prentice Hall.

Bushman, R. M., Indjejikian, R. J., & Smith, A. (1996). CEO Compensation: The role of individual performance evaluation. *Journal of Accounting and Economics, 21*(2), 161-193.

Cappelen, A. W., & Tungodden, B. (2012). Insentiver og innsats. *Magma - Tidsskrift for økonomi og ledelse*, 38-44.

Cawley, B. D., Keeping, L. M., & Levy, P. E. (1998). Participation in the performance appraisal process and employee reactions: A meta-analytic review of field investigations. *Journal of applied psychology, 83*(4), 615-633.

Cerasoli, C. P., Nicklin, J. M., & Ford, M. T. (2014). Intrinsic motivation and extrinsic incentives jointly predict performance: a 40- year meta- analysis. *Psychological bulletin, 140*(4), 980-1008.

Che, Y-K., & Yoo, S-W. (2001). Optimal Incentives for Teams. *The American Economic Review, 91*(3), 525-541.

Cleveland, J. N., Murphy, K. R., & Williams, R. E. (1989). Multiple uses of performance appraisal: Prevalence and correlates. *Journal of Applied Psychology, 74*(1), 130-135.

Colbjørnsen, T., Bragelien, I., Gjesdal, F., Hagen, K.P., Heum, P., & Salvanes, K.G (2000) *Resultatavhengig belønning: En utredning skrevet for Kreditkassen.* SNF-rapport nr. 25/00, Bergen.

Colquitt, J. A., Conlon, D. E., Wesson, M. J., Porter , C. O. L. H., & Ng, K. Y. (2001). Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology, 86*(3), 425-445.

Dahle, D. Y. (2014). *Orden og oppførsel: Karakterer på jobben?* Oslo: Gyldendal Norsk Forlag.

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology, 18*(1), 105-115.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior.* New York: Plenum.

Deci, E. L., Ryan, R. M., & Koestner, R. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125*(6), 627-668.

DeNisi, A. S., & Kluger, A. N. (2000). Feedback effectiveness: can 360-degree appraisals be improved? *The Academy of Management Executive, 14*(1), 129-139.

DeNisi, A. S., & Pritchard, R. D. (2006). Performance Appraisal, Performance Management and Improving Individual Performance: A Motivational Framework. *Management and Organization Review, 2*(2), 253-277.

Duffy, K. E., & Webber, R. E. (1974). On «relative» rating systems. *Personnel Psychology, 27*(2), 307-311.

Eichenwald, K. (2012, 24. juli). *Microsoft´s Lost Decade.* Hentet fra http://www.vanityfair.com/news/business/2012/08/microsoft-lost-mojo-steve-ballmer

Eisenhardt, K. (1989). Agency Theory: An Assessment and Review. *Academy of Management, 14*(1), 57-74.

Erez, M. (1977). Feedback: A necessary condition for the goal setting-performance relationship. *Journal of Applied Psychology, 62*(5), 624-627.

Feltham, G. A., & Xie, J. (1994). Performance measure congruity and diversity in multi-task principal-agent relations. *The Accounting Review, 69*(3), 429-453.

French, J. R. (1968). The conceptualization and measurement of mental health in terms of self-identity theory. I S. B. Sells, *The definition and measurement of mental health* (s. 136-159). Washington, D.C: Department of Health, Education and Welfare.

Frey, B., & Jegen, R. (2001). Motivation Crowding Theory: A Survey of Empirical Evidence. *Journal of Economic Surveys, 15*(5), 589-611.

Fulk, J., Brief, A. P., & Barr, S. H. (1985). Trust-in-supervisor and percieved fairness and accuracy in performance evaluations. *Journal of Business research, 13*(4), 301-313.

Gagné, M., & Deci, E. L. (2005). Self-determination theory and work motivation. *Journal of organizational behavior, 26*(4), 331-362.

Ghauri, P., & Grønhaug, K. (2010). *Research Methods in Business Studies* (4. utg.). Harlow: Pearson Education Limited.

Gibbons, R. (1987). Piece-rate incentive schemes. *Journal of Labor Economics, 5*(4), 413-429.

Gibbs, M., Merchant, K. A., Van der Stede, W. A., & Vargus, M. E. (2004). Determinants and Effects of Subjectivity in Incentives. *The Accounting Review, 79*(2), 409-436.

Gneezy, U., & Rustichini, A. (2000). Pay Enough Or Don't Pay At All. *The Quarterly Journal of Economics, 115*(3), 791-810.

Goffin, R. D., Gellatly, I. R., Paunonen, S. V., Jackson, D. N., & Meyer, J. P. (1996). Criterion validation of two approaches to performance appraisal: The behavioral observation scale and the relative percentile method. *Journal of Business and Psychology, 11*(1), 23-33.

Greenberg, J. (1990). Organizational justice: Yesterday, today, and tomorrow. *Journal of Management, 16*(2), 399-432.

Greenwald, B. C. (1986). Adverse Selection in the Labour Market. *The Review of Economic Studies, 53*(3), 325-347.

Guralnik, O., & Wardi, L. A. (2003). Forced distribution: A controversy. *Society for Human Resource Management White Paper Series.*

Guralnik, O., Rozmarin, E., & So, A. (2004). Forced distribution: Is it right for you? *Human Resource Development Quarterly, 15*(3), 339–345.

Hall, P. A., & Soskice, D. W. (2001). *Varieties of capitalism: the institutional foundations of comparative advantage.* Oxford: Oxford University Press.

Hayes, R. M., & Schaefer, S. (2000). Implicit contracts and the explanatory power of top executive compensation. *Rand Journal of Economics, 31*(2), 273-293.

Henderson, I. (2011). *Human Resource Management for MBA Students* (2. utg.). London: CIPD.

Hendrikse, G. W. (2003). *Economics and Management of Organisations: co-ordination, motivation and strategy* (2. utg.). London: McGraw-Hill Education.

Heneman, R. L. (1986). The relationship between supervisory ratings and results-oriented measures of performance: A meta-analysis. *Personnel Psychology, 39*(4), 811-826.

Henry, G. T. (1990). *Practical sampling.* Thousand Oaks: SAGE Publications.

Holmstrom, B. (1979). Moral Hazard and Observability. *The Bell Journal of Economics, 10*(1), 74-91.

Holmstrom, B. (1982). Moral hazard in teams. *The Bell Journal of Economics, 13*(2), 324-340.

Holmstrom, B. (1999). Managerial Incentive Problems: A Dynamic Perspective. *The Review of Economic Studies, 66*(1), 169-182.

Holmstrom, B., & Milgrom, P. (1991). Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics & Organization, 7*(1), 24-52.

Ittner, C. D., Larcker, D. F., & Meyer, M. W. (2003). Subjectivity and the weighting of performance measures: Evidence from a balanced scorecard. *The Accounting Review, 78*(3), 725-758.

Ittner, C. D., Larker, D. F., & Rajan, M. V. (1997). The Choice of Performance Measures in Annual Bonus Contracts. *The Accounting Review, 72*(2), 231-255.

Jacobsen, D. I. (2015). *Hvordan gjennomføre undersøkelser?: Innføring i samfunnsvitenskapelig metode* (3. utg.). Oslo: Cappelen Damm Akademisk.

Jawahar, I. M., & Williams, C. R. (1997). Where all the children are above average: The performance appraisal purpose effect. *Personnel Psychology, 50*(4), 905–925.

Jenkins, G. D., Mitra, A., Gupta, N., & Shaw, J. D. (1998). Are Financial Incentives Related to Performance? A Meta-Analytic Review of Empirical Research. *Journal of Applied Psychology, 83*(5), 777-787.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure. *Journal of Financial Economics, 3*(4), 305-360.

Johannessen, A., Kristoffersen, L., & Tufte, P. A. (2011). *Forskningsmetode for økonomisk-administrative fag* (3. utg.). Oslo: Abstrakt Forlag.

Johannessen, A., Tufte, P. A., & Kristoffersen, L. (2006). *Introduksjon til samfunnsvitenskapelig metode* (3. utg.). Oslo: Abstrakt forlag.

Jones, B., & Rock, D. (2015). *Why more and more companies are ditching performance ratings*. Hentet fra https://hbr.org/2015/09/why-more-and-more-companies-are-ditching-performance-ratings

Judge, T. A, & Ferris, G. R. (1993). Social Context of Performance Evaluation Decisions. *Academy of Management Journal, 36*(1), 80-105.

Kandel, E., & Lazear, E. P. (1992). Peer Pressure and Partnerships. *Journal of Political Economy, 100*(4), 801-817.

Kaplan, R. S., & Norton, D. P. (1996). *The Balanced Scorecard: Translating Strategy into Action.* Boston, MA: Harvard Business School Press.

Kaufmann, G., & Kaufmann, A. (2009). *Psykologi i organisasjon og ledelse.* Bergen: Fagbokforlaget.

Kuvaas, B. (2005). Belønning og motivasjon: ytre og indre motivasjon som kilder til innsats og kvalitet i arbeidslivet. I K. Knudsen, & A. Ryan, *Hvordan kan frynsegoder bli belønning?* (s. 29-51). Oslo: Cappelen Damm Akademisk.

Kuvaas, B. (2008). *Lønnsomhet gjennom menneskelige ressurser: evidensbasert HRM.* Bergen: Fagbokforlaget.

Kuvaas, B., & Dysvik, A. (2012). *Lønnsomhet gjennom menneskelige ressurser: Evidensbasert HRM* (2. utg.). Bergen: Fagbokforlaget.

Kvale, S., & Brinkmann, S. (2009). *Den kvalitativa forskningsintervjun* (2. utg.). Lund: Studentlitteratur.

Landy, F. J., & Farr, J. L. (1980). Performance rating. *Psychological Bulletin, 87*(1), 72-107.

Langeland, O. (1999). *Kompensasjon og motivasjon: overskuddsdeling og medeierskap i norske bedrifter.* Doktoravhandling. Nr. 315, Fafo, Oslo.

Latham, G. P., & Locke, E. A. (1979). Goal Setting - A Motivational Technique That Works. *Organizational Dynamics, 8*(2), 68-80.

Lawler, E. E. (2002). The Folly of Forced Ranking. *Strategy & Business, 28*, 28-32.

Lazear, E. P. (1989). Pay Equality and Industrial Politics. *Journal of Political Economy, 97*(3), 561-580.

Lazear, E. P. (1998). *Personnels Economics for Managers.* New York: John Wiley & Sons Inc.

Lazear, E. P. (2000). Performance pay and productivity. *The American Economic Review, 90*(5), 1346-1361.

Lazear, E. P, & Gibbs, M. (2015). *Personell Economics in Practice* (3. utg.). Hoboken: John Wiley & Sons.

Leventhal, G. S. (1980). What should be done with equity theory? New approaches to the study of fairness in social relationship. I K. Gergen, M. Greenberg, & R. Willis, *Social exchange: Advances in theory and research* (s. 27-55). New York: Plenum Press.

Levy, P. E., & Williams, J. R. (2004). The Social Context of Performance Appraisal: A Review and Framework for the Future. *Journal of Management, 30*(6), 881-905.

Locke, E. A., & Latham, G. P. (2002). Building a Practically Useful Theory of Goal Setting and Task Motivation. *American Psychologist, 57*(9), 705-717.

Lunde, H., & Grini, K. H. (2007). Bonus – hvor mye og til hvem? *Statistisk sentralbyrå,* rapport nr. 18.

Macdougall, N. (1991). The story behind salary increases. *CMA - the Management Accounting Magazine, 65*, 34.

Maddux, R. (1997). *Effective performance appraisals.* London: Kogan Page.

Martén, I., & Whittaker, P. (2015). *Lower, and More Volatile, Oil Prices: What They Mean and How to Respond*. Hentet fra https://www.bcgperspectives.com/content/articles/energy_environment_lower_more_volatile_oil_prices/

McAllister, D. J. (1995). Affect- and Cognition-Based Trust as Foundation for Interpersonal Cooperation in Organizations. *Academy of Management Journal, 38*(1), 24-59.

McBriarty, M. A. (1988). Performance appraisal: Some unintended consequences. *Public Personnel Management, 17*(4), 421–434.

Merchant, K. A. (1989). *Rewarding Results: Motivating Profit Center Managers.* Boston, MA: Harvard Business School Press.

Merchant, K. A., & Manzoni, J. F. (1989). The achievability of budget targets in profit centers: A field study. *The Accounting Review, 64*, 539-558.

Milgrom, P. (1988). Employment Contracts, Influence Activity and Efficient Organization. *The Journal of Political Economy, 96*(1), 42-60.

Milgrom, P. (1992). *Economics, organization and management.* Englewood Cliffs, N.J: Prentice-Hall.

Milgrom, P., & Roberts, J. (1988). An Economic Approach to Influence in Organizations. *American Journal of Sociology, 94,* 154-179.

Murphy, K. R. (2008). Explaining the weak relationship between job performance and ratings of job performance. *Industrial and organizational psychology, 1*(2), 148-160.

Murphy, K. R., & Cleveland, J. N. (1991). *Performance Appraisal: An Organizational Perspective.* Boston: Allyn & Bacon.

Murphy, K. R., & Cleveland, J. N. (1995). *Understanding Performance Appraisal: Social, Organizational and Goal-Based Perspectives.* Thousand Oaks: SAGE Publications.

Murphy, K. R., & Oyer, P. (2003). Discretion in executive incentive contracts: Theory and evidence. *University of Southern California and Stanford Univerity Working paper.*

Murphy, K. R., Jako, R. A., & Anhalt, R. L. (1993). Nature and Consequences of Halo Error: A Critical Analysis. *Journal of Applied Psychology, 78*(2), 218-225.

Nathan, B. R., & Alexander, R. A. (1988). A comparison of criteria for test validation: A metal-analytic investigation. *Personnel Psychology, 41*(3), 517-535.

Nordhaug, O. (2002). *LMR - Ledelse av menneskelige ressurser: Målrettet personal- og kompetanseledelse* (3. utg.). Oslo: Universitetsforlaget.

Olson, C. A., & Davis, G. M. (2003, Mars). *Pros and cons of forced ranking and other relative performance ranking systems*. Hentet fra Society for Human Resource Management: http://www.shrm.org/hrresources/lrpt_published /CMS_003991

Pennings, J. M., & Woiceshyn, J. (1987). A typology of organizational control and its metaphors. *Research in the Sociology of Organizations, 5,* 73-104.

Pfeffer, J. (1998). *The human equation: Building profits by putting people first.* Boston: Harvard Business School.

Pfeffer, J., & Sutton, R. I. (2006). Evidence-based management. *Harvard Business Review, 84*(1), 62–74.

Prendergast, C. (1999). The Provision of Incentives in Firms. *Journal of Economic Literature, 37*(1), 7-63.

Prendergast, C., & Topel, R. (1993). Discretion and bias in performance evaluation. *European Economic Review*, *37*(2), 355-365.

Ross, S. A. (1973). The Economic Theory of Agency: The Principal's Problem. *The American Economic Review, 63*(2), 134-139.

Ryan, R. M., & Deci, E. L. (2000). Self- Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist, 55*(1), 68-78.

Salamin, A., & Hom, P. W. (2005). In search of the elusive U-shaped performance-turnover relationship: Are high performing Swiss bankers more liable to quit? *Journal of Applied Psychology, 90*(6), 1204-1216.

Saal, F. E., Downey, R. G., & Lahey, M. A. (1980). Rating the Ratings: Assessing the Psychometric Quality of Rating Data. *Psychological Bulletin, 88*(2), 413–428.

Saunders, M. N. K., Lewis, P., & Thornhill, A. (2012). *Research methods for business students* (6. utg.). Harlow: Pearson.

Schleicher, D. J., Bull, R. A., & Green, S. G. (2009). Rater Reactions to Forced Distribution Rating Systems. *Journal of Management, 35*(4), 899-927.

Schrage, M. (2000, Februar 21). How the bell curve cheats you. *Fortune, 141*(4), 296.

Skinner, B. (1953). *Science and human behavior.* New York: Macmillan.

Spence, J. R., & Keeping, L. M. (2010). The impact of non-performance information on ratings of job performance: A policy-capturing approach. *Journal of organizational behavior, 31*(4), 587-608.

Stewart, S. M., Gruys, M. L., & Storm, M. (2010). Forced distribution performance evaluation systems: Advantages, disadvantages and keys to implementation. *Journal of Management & Organization, 16*(1), 168-179.

Sutton, A. W., Baldwin, S. P., Wood, L., & Hoffman, B. J. (2013). A Meta-Analysis of the Relationship Between Rater Liking and Performance Ratings. *Human Performance, 26*(5), 409-429.

Thorndike, E. L. (1920). A Constant Error in Psychological Ratings. *Journal of Applied Psychology, 4*(1), 25-29.

Tirole, J. (1992). Collusion and the Theory of Organizations. I J. Laffont, *Advances in Economic Theory: Sixth World Congress* (s. 151-206). Cambridge: Cambridge University Press.

Truby, M. (2001, 29. april). Age-bias claims jolt Ford culture change. *The Detroit News*.

Vance, A., & Davidhizar, R. (1998). Motivating the Minimal Performer. *Hospital Topics, 76*(4), 8-12.

Wagner, S. H., & Goffin, R. D. (1997). Differences in accuracy of absolute and comparative performance appraisal methods. *Organizational Behavior and Human Decision Processes, 70*(2), 95-103.

Weibel, A., Rost, K., & Osterloh, M. (2007). Crowding-Out of Intrinsic Motivation: Opening the Black Box. *SSRN Working Paper*.

Werther, W., & Davis, K. (1989). *Human Resources and Personnel Management.* New York: McGraw-Hill.

Williams, M. (2004). Generalization/Generalizability in Qualitative Research. I M. Lewis-Beck, A. Bryman, & T. Liao, *The Sage Encyclopedia of Social Science Research Methods* (s. 420-421). Thousand Oaks: SAGE Publications.

Yin, R. K. (2009). *Case Study Research: Design and Methods* (4. utg.). Thousand Oaks: SAGE Publications.

Yin, R. K. (2014). *Case Study Research: Design and Methods* (5. utg.). Thousand Oaks: SAGE Publications.

Zenger, T. R. (1992). Why do employers only reward extreme performance? Examining the relationships among performance, pay, and turnover. *Administrative Science Quarterly, 37*(2), 198-219.

9. Vedlegg

9.1 Intervjuguide

I det følgende presenteres intervjuguidene vi benyttet i bedriftene. Vi hadde en intervjuguide til ledere med evalueringsansvar og en intervjuguide til ansatte uten evalueringsansvar. Vi gjør oppmerksom på at intervjuguiden ikke er uttømmende, da det ble stilt mange oppfølgingsspørsmål i samtlige intervjuer. I tillegg ble intervjuguidene tilpasset den enkelte bedrift ut fra informasjonen vi hadde på forhånd.

9.1.1 Informasjon som gis i starten av intervjuet

- Takke for at respondenten tok seg tid til å komme.
- Presentere oss selv.
- Forklare respondenten om oppgavens formål. Fortell hva vi legger i skjønnsmessige vurderinger, karakterer, belønningssystem og andre viktige begreper.
- Understreke at vi er helt uavhengige og ikke har en politisk agenda.
- Informere om at respondenten og bedriften vil bli behandlet anonymt.
- Informere om at respondenten ikke trenger å snakke om noe vedkommende ikke ønsker. Dersom en ikke ønsker å besvare ett spørsmål underveis eller ønsker å trekke seg fra intervjuet, blir dette respektert.
- Informere om at samtalen blir tatt opp på bånd, noe som gjøres av praktiske årsaker. Det er for at vi skal kunne slippe å ta notater og i stedet fokusere på å få til en god dialog. Dersom respondenten ikke ønsker å bli tatt opp på bånd, vil vi ta hensyn til det. Opptakene slettes umiddelbart etter transkribering, og transkriberingen kan sendes til intervjuobjekter som ønsker det.
- Bekrefte avtalt tidsbruk på intervjuet (1-2 timer)

9.1.2 Spørsmål til ledere med evalueringsansvar

Dette er spørsmål til ledere med oversikt over utformingen av belønningssystemet og systemer for prestasjonsevalueringer. Hensikten er å undersøke formålet med og virkninger av systemene.

Introduksjon

- Kan du fortelle om din bakgrunn og stilling i bedriften?

Belønningssystemet

- Kan du fortelle om belønningssystemet i bedriften?
    - Fast lønn vs. prestasjonsbasert lønn
        - Hvor ofte og hvor store er eventuelle bonuser/lønnsøkninger for ikke-ledende ansatte?
    - Finansiell vs. ikke-finansiell belønning
- Hva er formålet med belønningssystemet?
- Hvordan tror du belønningssystemet påvirker bedriftens evne til å tiltrekke og beholde ønsket arbeidskraft?
    - Vet du om noen som har sluttet grunnet misnøye med belønningssystemet?
- Hvordan synes du dagens belønningssystem virker?
    - Er det noe du ville endret?
- Hvordan har belønningssystemet blitt endret som følge av nedgangstidene?

Prestasjonsevalueringer

- Hvilke kriterier legges til grunn ved fastsettelse av den enkeltes belønning?
    - Subjektive vs. objektive måltall
    - Relative eller absolutte prestasjonsmål
- Hvordan og hvor ofte evalueres ansattes prestasjoner?
- Hvor mange er med i evalueringen av den enkelte?
- Hvor mange ansatte har du evalueringsansvar for?
    - Hvor store/små er gruppene ledere evaluerer?
- Justeres krav til den enkelte over tid?
- Hvordan kommuniseres det til ansatte hvilke kriterier de måles etter?
- Hvordan og hvor ofte gis de ansatte tilbakemeldinger?

- Hvordan opplever du at bruk av leders skjønn i vurderingene påvirker forholdet mellom leder og ansatt?
    - Eventuelle konflikter som følge av karaktergivning
    - Har du opplevd at ansatte prøver å smiske/påvirke deg for å få bedre evaluering og belønning?
    - Hva synes du om å gi ansatte dårlige karakterer?
- Hvilke fordeler og ulemper ser du med bruk av skjønn i evalueringene?

Karakterer og tvungen fordeling

- Hvordan er systemet for karakterer bygget opp?
- Foreligger krav eller forventning til fordeling av ansattes karakterer?
    - Eksempelvis normalfordeling, rangering fra beste til dårligste medarbeider, gjennomsnittskarakter eller fastsatte prosentandeler
- Kan du fortelle om kalibreringsprosessen? (Hvordan dere kommer frem til endelig karakterfordeling?)
- Hvor tydelig skiller du mellom gode og dårlige prestasjoner?
    - Bruker du hele karakterskalaen?
- Hvordan tror du karakterene påvirker ansattes motivasjon?
- Hvordan tror du karakterene med tilhørende belønninger påvirker turnover?
- Hvordan tror du karaktergivningen med tilhørende belønninger påvirker kulturen i bedriften?
- Hvordan opplever du at bruk av skjønn og karakterer påvirker insentiver til samarbeid?
    - Opplever du at ansatte samarbeider godt?
- Hvilke fordeler og ulemper ser du ved å gi ansatte karakterer?
- Hvilke fordeler og ulemper ser du ved bruk av normalfordeling/fastsatte prosentandeler/rangering/gjennomsnittskarakter?

Oppfatninger om ikke-ledende ansatte

- Hvordan tror du belønningssystemet/karaktersystemet virker inn på de ansattes prioritering mellom ulike arbeidsoppgaver?
    - Er det enkelte oppgaver som ofte neglisjeres hos dere?
- Har de ansatte klart definerte arbeidsoppgaver?
    - Komplekse eller spesifikke oppgaver

  - Mange eller få
- I hvilken grad opplever du at ansatte har kontroll over sine resultater?
  - Er det ukontrollerbare forhold som påvirker? (støy, markedsforhold, oljepris, osv)
  - Hvordan tas dette hensyn til i vurderingene?
- Hvilken holdning tror du de ansatte har til belønningssystemet og at det brukes skjønn og karakterer i prestasjonsevalueringene?
  - Hvem er mest/minst fornøyde?
  - Tror du de opplever det rettferdig/urettferdig?
  - Opplever du at dine ansatte har tillit til din evaluering av deres prestasjoner?

Avslutningsvis

- Er det noe du vil spørre oss om?
- Få kontaktinformasjonen til vedkommende.
- Opplyse om at transkriberingen kan tilsendes informanten for godkjenning dersom det er ønskelig.
- Takke respondenten for at vedkommende tok seg tid til å bli intervjuet.

9.1.3 Spørsmål til ikke-ledende ansatte uten evalueringsansvar

Dette er spørsmål til ansatte i bedriften uten evalueringsansvar. Hensikten er å undersøke oppfatninger og virkninger av bedriftens belønningssystem og system for prestasjonsevalueringer.

Introduksjon

- Kan du fortelle om din bakgrunn og din stilling i bedriften?
- Har du klart definerte arbeidsoppgaver?
  - Komplekse eller spesifikke
  - Mange eller få
- Hvor delaktig er du i prosessen med å fastsette mål som angår deg?
- I hvilken grad har du kontroll over dine resultater?
  - Er det mange ukontrollerbare forhold som påvirker prestasjonene?

Belønningssystem

- Hvilke typer belønninger opplever du i din arbeidshverdag?
    - Finansielle og ikke-finansielle
- Hva slags form for belønning er viktig for deg?
    - Opplever du at disse er godt ivaretatt i bedriften?
- I hvilken grad opplever du belønningene som en driver for motivasjon?
- Hvordan synes du dagens belønningssystem virker?
    - Er det noe du ville endret?

Prestasjonsevalueringer

- Hvilke kriterier evalueres du etter?
- I hvilken grad opplever du at kriteriene dekker alle relevante aspekter ved din arbeidshverdag?
- Hvor mange er med i evalueringen av dine prestasjoner?
- I hvilken grad opplever du at din leder har mulighet til å observere din innsats og dine prestasjoner?
- Hvordan og hvor ofte får du tilbakemeldinger relatert til måloppnåelse?
    - Er lederne like gode på positive og kritiske tilbakemeldinger?
- Hvordan opplever du at skjønnsmessige vurderinger påvirker forholdet mellom leder og ansatt?
    - Har du opplevd at enkelte har fått dårligere relasjon etter karaktergivning?
    - Har du opplevd at noen prøver å påvirke leder for å få bedre belønning/evaluering?
- Har du tillit til at din leder foretar informerte og objektive prestasjonsevalueringer?
    - Hva påvirker din tillit til leder?
    - Har du opplevd at trynefaktor har påvirket prestasjonsevalueringene?
- Opplever du at ledere skiller tydelig mellom sterke og svake prestasjoner?
    - Brukes hele karakterskalaen?
- Hvilke fordeler og ulemper ser du med bruk av skjønn i prestasjonsevalueringer?

Karakterer og tvungen fordeling

- I hvilken grad motiveres du av medarbeidersamtalen og karakteren du får?
- Fokuserer du mest på karakteren eller tilbakemeldingene du får?
- Hvordan påvirker det deg om du får en svakere karakter enn forventet?

- Sammenligner du dine karakterer med andre ansattes karakterer?
- Hvordan opplever du at karakterene og belønningssystemet påvirker bedriftskulturen?
    - Opplever du at ansatte samarbeider godt seg i mellom?
- Hvor stor sammenheng ser du mellom innsatsen du legger ned og karakteren/ belønningen du mottar?
    - Hvor ofte opplever du eventuelle avvik?
    - Hva tror du eventuelle avvik skyldes?
- Hvilke fordeler og ulemper ser du med å gi ansatte karakterer?
- Hvilke fordeler og ulemper ser du med forventningen om en gitt karakterfordeling?
- Hvordan tror du karakterer og belønningssystemet påvirker grad av turnover?
    - Vet du om noen som har startet på grunn av belønningene?
    - Vet du om noen som har sluttet på grunn av misnøye med karakter eller belønning?

Avslutningsvis

- Er det noe du vil legge til?
- Er det noe du vil spørre oss om?
- Få kontaktinformasjonen til vedkommende.
- Opplyse om at transkriberingen kan tilsendes informanten for godkjenning dersom det er ønskelig.
- Takke respondenten for at vedkommende tok seg tid til å bli intervjuet.