

zuordenbar ist, wie dies bspw. bei der Erfassung von organisationalen Daten über Führungskräftebefragungen der Fall ist (vgl. Enticott 2004: 744f.). Ohnehin scheint die hierarchische Position der Befragten positiv mit deren Leistungseinschätzung zu korrelieren (Light 2006: 16): „There is other evidence, that the view of performance looks better from the top [...]“. Während für Führungskräfte angenommen wird, dass sie ein starkes Eigeninteresse an einer positiven Darstellung der Behördenleistung haben, wird Vollzugspersonal eine eher zynische, aber teilweise auch fragmentierte Sicht zugesprochen (vgl. Enticott 2004: 745). Zusammenfassend legen Enticott et al. (2009: 246-247) nahe, dass Spitzenbeamte eine weniger zuverlässige Informationsquelle darstellen, als *front-line officers*.

Für die vorliegende Studie wurde gezielt die Mitarbeiterebene adressiert, um das bei Führungskräftebefragungen gegebene Risiko geringer Validität und Reliabilität zu umgehen. Durch die fragebogenbasierte Erhebung und die – insbesondere durch die (teilweise) Mittlerrolle von Gewerkschaft und Fachverband – glaubhafte Versicherung ihrer Anonymität sollte die Problematik einer Verzerrung durch angepasstes Verhalten, wie auch Reliabilitätsprobleme durch zu hohes Gewicht einzelner Befragter, erfolgreich minimiert worden sein. Darüber hinaus sollte aufgrund des gewählten Untersuchungsfeldes der *Gewerbeaufsichtsverwaltung* die Problematik einer starken Fragmentierung der Mitarbeiterperspektive ebenfalls weitgehend vermieden worden sein: Die Vollzugsmitarbeiter arbeiten überwiegend eigenverantwortlich mit umfassender Kompetenz an Fällen, die in ihr Spezialgebiet fallen. Eine hochgradige Arbeitsteilung wie sie bspw. im Bereich der Leistungsverwaltung zu beobachten ist, die zu einem "Tunnelblick" der Mitarbeiter führen kann, ist hier auszuschließen. Durch die überwiegend noch vorhandene, auf eine technische Ausbildung aufsetzende Fachausbildung des Genehmigungs- und Aufsichtspersonals kann darüber hinaus auch von einer einheitlichen professionellen Sichtweise und Ethik der Akteure ausgegangen werden, die auf einer anerkannten *best practice* basiert.

5.5.4. Common Source Bias

An die Diskussion zur Verwendung von Perzeptionsdaten schließt sich unmittelbar die sich aus der Anlage der Untersuchung ergebende Frage nach einem möglichen *Common Source Bias* im erhobenen Datensatz an. *Common Source Bias* kann als Sonderform des *Common Method Bias* gesehen werden, welcher einen systematischen Zusammenhang zwischen Konstrukten allein aufgrund der Anwendung derselben Erhebungsmethoden thematisiert (vgl. Podsakoff et al. 2003: 879). *Common Source Bias* kann auftreten, wenn unabhängige und abhängige Variablen über Perzeptionsdaten ein und derselben Person erfasst werden (Podsakoff et al. 2003: 881). Dies ist für die vorliegende Untersuchung für die Vari-

ablen der *Führungskultur* und jene der abhängigen Variablen der Fall. Mögliche Ursachen für einen systematischen Zusammenhang zwischen den einzelnen Aussagen der Zielpersonen sind vielfältig (Podsakoff et al. 2003: 881): Antwortende Personen versuchen bekanntermaßen in sich konsistente Antworten zu geben, lassen sich bei ihren Antworten von Vermutungen über Zusammenhänge zwischen den Items leiten, versuchen – wie oben erwähnt – Kriterien sozialer Gewünschtheit zu entsprechen. Auch neigen sie dazu, Personen ihres Umfelds positiv zu bewerten und Antworten tendenziell zu bejahen. Schließlich spielen auch individuelle Stimmungen über alle Konstrukte hinweg eine Rolle.

Das Problem liegt nun im Gegensatz zu den oben beschriebenen Elitebefragungen nicht in der systematischen Überschätzung der Leistung, sondern in der Korrelation der Fehlerterme der berichteten unabhängigen und abhängigen Variablen. Hierdurch kann die Beziehung zwischen diesen beiden Konstrukten je nach Richtung des jeweiligen Zusammenhangs zwischen den Items über- oder unterschätzt werden (Meier/O'Toole 2010: 6). Dieses Problem ist in der einschlägigen Literatur anerkannt, wird jedoch anhand empirischer Befunde hinsichtlich seiner Stärke oft relativiert (Moynihan/Pandey 2005: 428, FN 1):

„The consensus is that while common-source bias does exist, its effect is not to invalidate relationships uncovered in studies employing self-reported data but to perhaps marginally attenuate the strength of the findings.“

Dennoch stellt *Common Source Bias* ein ernsthaftes Problem für die Validität auf Befragungsdaten basierender Befunde dar und sollte, wenn möglich, minimiert werden (vgl. Meier/O'Toole 2013). Der simpelste (und einzig wirklich effektive) Ansatz hierfür läge in getrennten Quellen für unabhängige und abhängige Variablen. Dies war bei der vorliegenden Untersuchung nicht möglich. Da die Berücksichtigung der *weichen* Performanzaspekte zentral für eine holistische Einschätzung administrativer Performanz ist, kann auf sie nicht verzichtet werden. Eine Differenzierung in verschiedene Akteursgruppen, mit deren Hilfe unabhängige und abhängige Variablen getrennt erhoben werden könnte, ist bei der adressierten Zielgruppe des Vollzugspersonals nicht möglich. Abgesehen vom fragwürdigen Erfolg dieses Ansatzes (Meier/O'Toole 2010: 20f.) ist dies aufgrund der gebotenen Anonymität der Befragung und der individuellen, kleingliedrigen Struktur der Verwaltungseinheiten nicht valide durchführbar. Vorgeschlagene statistische Verfahren (vgl. Podsakoff et al. 2003) setzen stets bei der Kovarianz der einzelnen Fragen an und erscheinen deshalb kein gangbarer Weg zur Behebung des beschriebenen Problems (Meier/O'Toole 2010: 21f.). Entsprechend bleibt nur, die berichteten Ergebnisse kritisch zu bewerten und vorsichtig zu interpretieren.

5.5.5. Aggregation von Befragungsdaten

Eine letzte grundsätzliche methodische Frage betrifft die Zulässigkeit der Aggregation von Befragungsdaten und insb. die Interpretation von Mittelwerten dieser Daten. Diese Verfahren werden bei den in Kapitel 6 vorgenommenen deskriptiven und ranggruppenvergleichenden Analysen verwandt. Ein solches Vorgehen setzt zwei Annahmen voraus: Dass erstens die Angehörigen unterschiedlicher Einheiten dasselbe Verständnis der Inhalte haben und dieselben Maßstäbe zu ihrer Bewertung anlegen und dass zweitens eine grundsätzliche Übereinkunft hinsichtlich der Einschätzung zwischen den befragten Akteuren vorliegt. Wird gegen diese Annahmen verstoßen, so werden Unterschiede innerhalb der Gruppen durch dieses Vorgehen verschleiert (West 2001: 10925). Zu welchem Grad die erste Annahme erfüllt werden kann, hängt vermutlich stark mit der Sozialisierung der Befragten innerhalb ihrer Behörde ab. Diese kann insofern als unbekannter Mediator der Einschätzungen gelten. Das Forschungsdesign der vorliegenden Untersuchung versucht diesem Problem durch die Fokussierung auf eine homogene Zielgruppe zu begegnen und so Verzerrungen durch Kognitionsunterschiede zumindest im Verhältnis zu fachgebietsübergreifenden oder internationalen Vergleichen auf ein vernachlässigbares Niveau zu reduzieren (vgl. Kapitel 5.5.3). Für die zweite Annahme kann bei der gewählten Befragungsmethodik nicht kontrolliert werden, so dass die Interpretation der entsprechenden Aggregatdaten mit Vorbehalt erfolgen sollte. Dort wo sinnvoll machen die berichteten Boxplots die Verteilung der Werte transparent.