

Diverging Stacked Bar Charts für Likert Skalen

Jörg Sellmann

joerg.sellmann@julius-kuehn.de

Zusammenfassung



Diverging Stacked Bar Charts werden häufig verwendet, um die Ergebnisse von Umfragen anzuzeigen, die auf der Basis von Likert-Skalen erhoben wurden.

Spezielle Programme als auch Excel, R und Python bieten die Möglichkeit, diese Grafiken zu erstellen. Hier wird gezeigt, wie mittels SAS-Base-Mitteln (DATA-Step, sgplot, Macro) eine solche Grafik zu gestalten ist.

Ein zweites Ziel dieses Beitrages ist, verschiedene Möglichkeiten der Programmsteuerung unter der Nutzung von Makrovariablen zu demonstrieren.

Likert-Skala



Wikipedia: Die Likert-Skala (nach Rensis Likert) ist ein Verfahren zur Messung persönlicher Einstellungen.

Die Skalen bestehen aus mehreren Items vom Likert-Typ. Diese sind Aussagen, denen die Befragten auf einer vorgegebenen mehrstufigen Antwortskala mehr oder weniger stark zustimmen oder die sie ablehnen können.

Bsp:

	Voll und ganz zufrieden	Weder zufrieden noch unzufrieden	Eher unzufrieden	Überhaupt nicht zufrieden	Eher zufrieden
Wie zufrieden waren Sie mit unserem Kundenservice?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Ausgangssituation



likert7.xlsx - Excel

Start | Einfügen | Seitenlayout | Formeln | Daten | Überprüfen | Ansicht | JMP | ACROBAT | Was möchten Sie tun?

Sellmann, Jörg | Freigeben

	A	B	C	D	E	F	G	H
1		Komplett unzufrieden	Meist unzufrieden	Etwas unzufrieden	weder zufrieden noch unzufrieden	Etwas zufrieden	Meist zufrieden	Rundum zufrieden
2	Frage 1	0	8	34	23	50	28	7
3	Frage 2	4	12	28	30	61	12	3
4	Frage 3	7	10	47	26	37	17	6
5	Frage 4	3	23	33	20	40	23	8
6	Frage 5	8	18	20	30	55	10	9
7	Frage 6	2	14	25	45	39	15	10
8	Frage 7	1	5	8	41	90	4	1
9	Frage 8	5	9	17	27	80	8	4

Bereit | 120 %

Ausgangssituation



Excel screenshot showing the initial data matrix in the 'Daten' sheet. The data is as follows:

	A	B	C	D	E
1	Ursache	Gar nicht	Weniger	Eher	Immer
2	Unterbrechungen	3	10	13	5
3	Zeitdruck	1	9	12	9
4	schlechtere Qualität	3	10	14	4
5	Arbeit anderer	5	15	9	2
6					

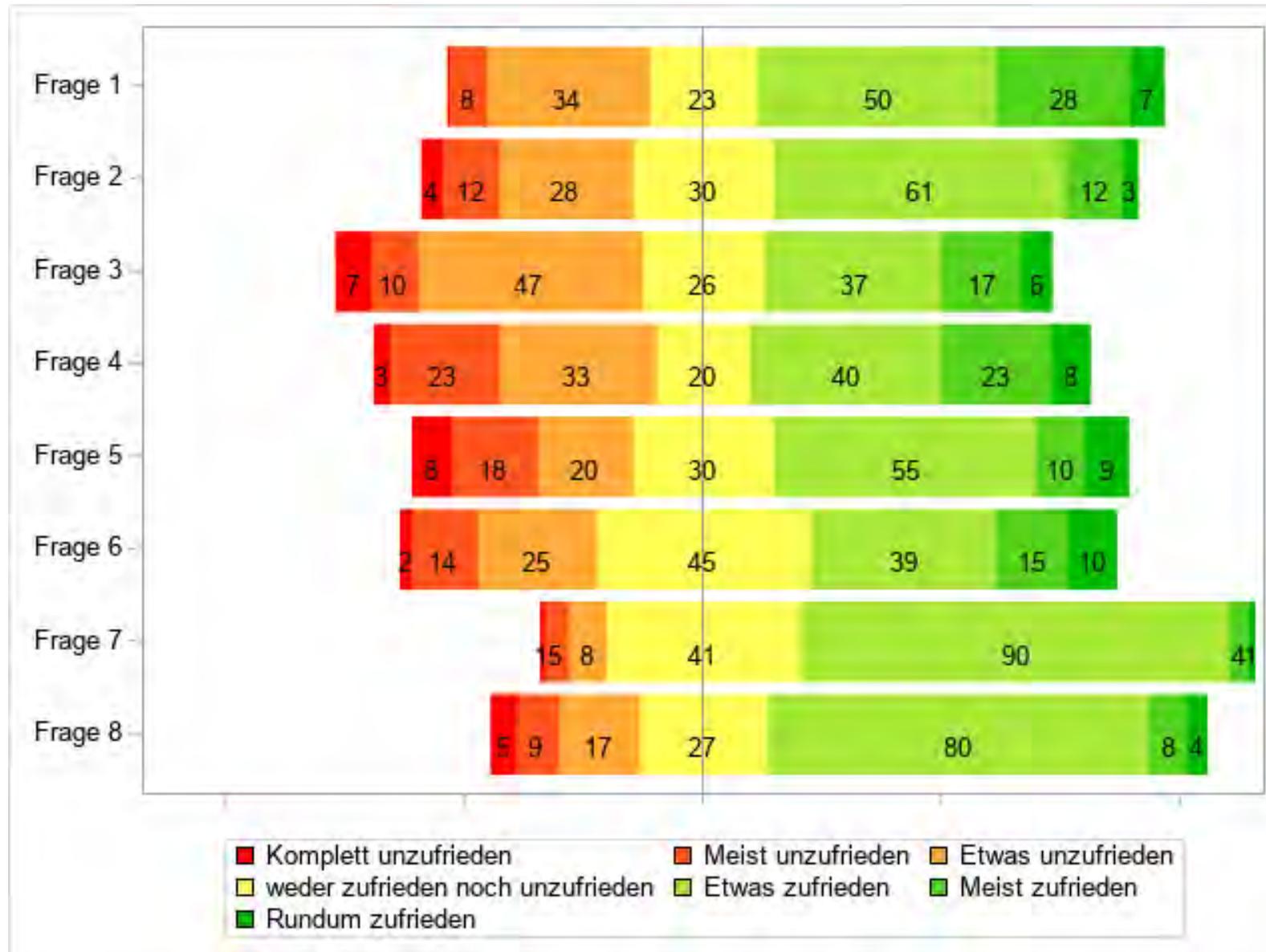
Datenmatrix

„SAS konforme“ Kodierung

Excel screenshot showing the SAS-compliant coding in the 'Code' sheet. The data is as follows:

	A	B	C	D	E	F
1	ID	Value	Rang	Faktor	Fillcolor	
2	Grad	Gar nicht	1	-1	lightred	
3	Grad	Weniger	2	-1	Gold	
4	Grad	Eher	3	1	lightblue	
5	Grad	Immer	4	1	lightgreen	
6						

Endergebnis

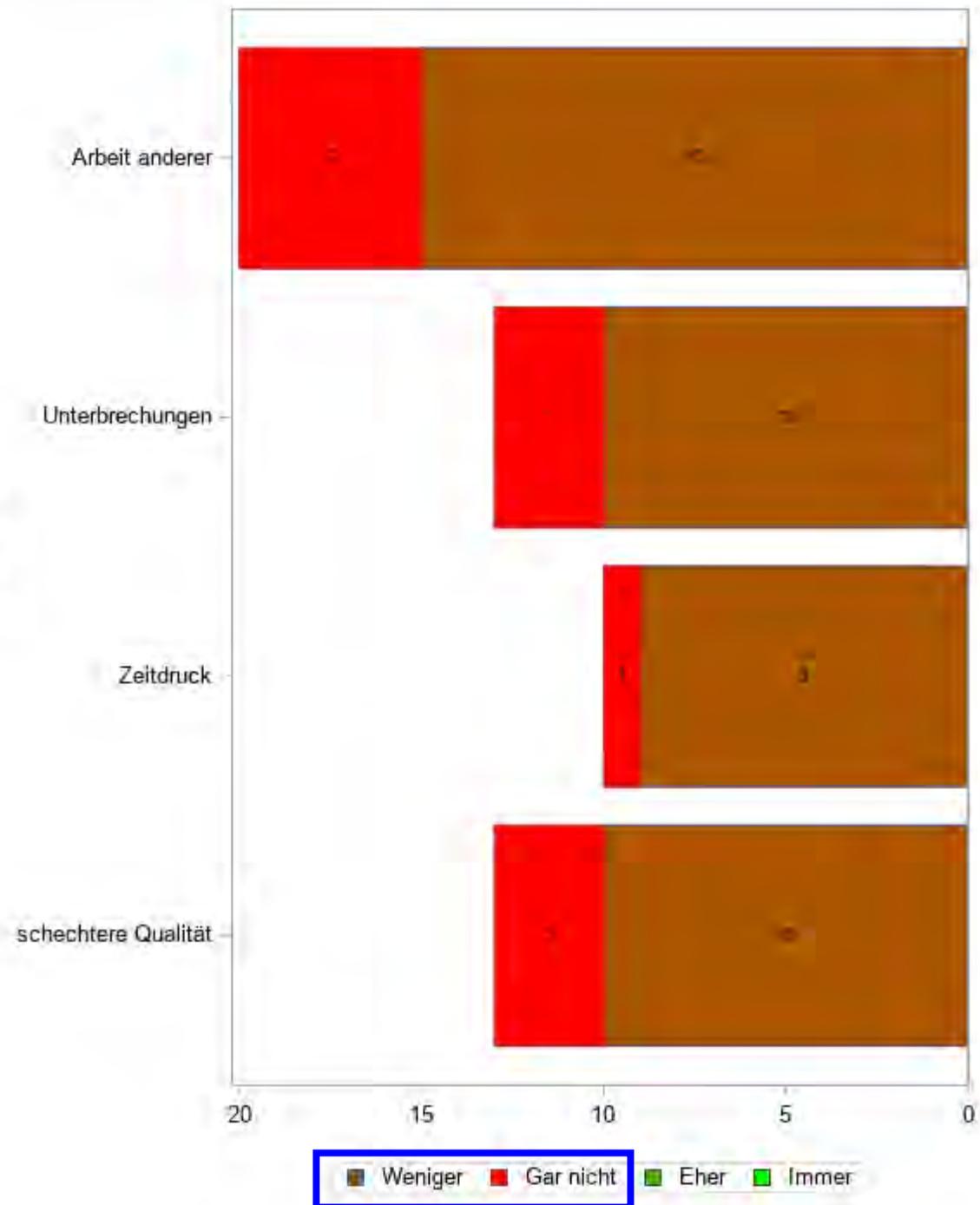


Wo liegt das eigentliche Problem?

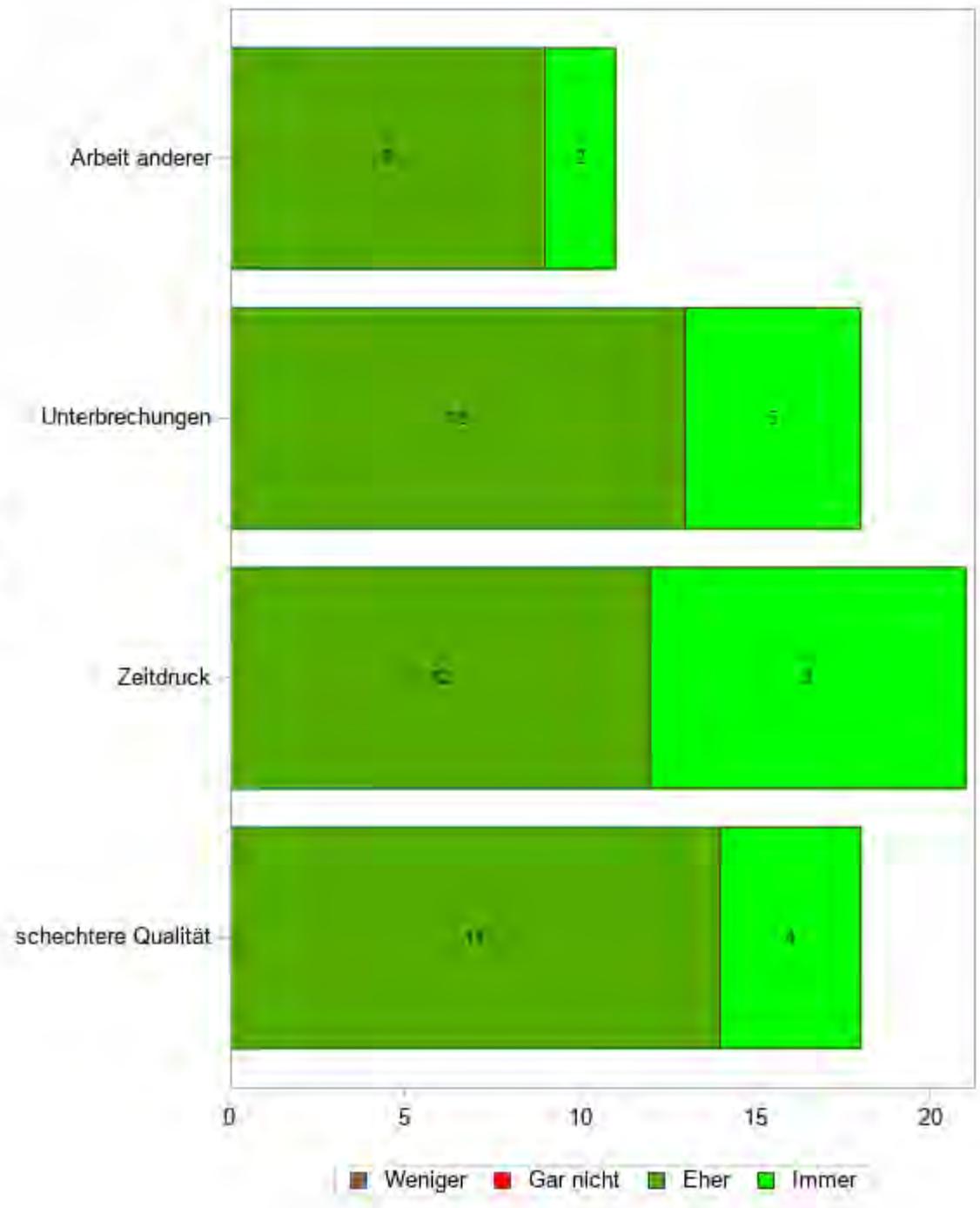


- SAS kennt doch Stacked Bars
- Wir brauchen aber Stacked Bars gleichzeitig in 2 Richtungen
- Wenn die Anzahl der Kategorien ungerade ist, müssen wir die mittlere Kategorie irgendwie aufteilen
- Bei Standard-Segmentlabeln haben wir dann aber 2 Werte in der Grafik für die mittlere Kategorie
- Im folgenden 6 Versuche, das Problem „einfach“ zu lösen, dabei wird nur von einer geraden Zahl von Kategorien ausgegangen

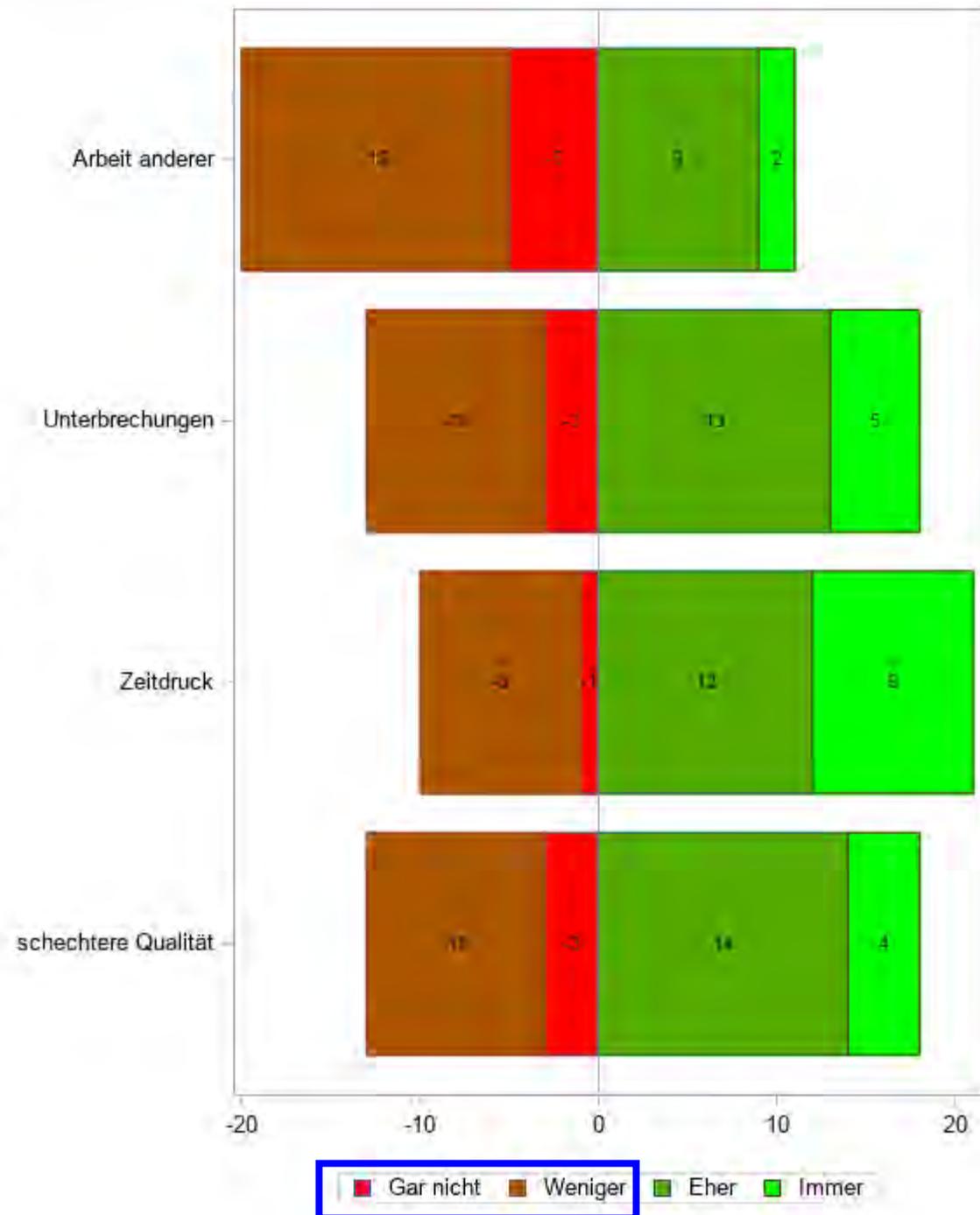
Ursache



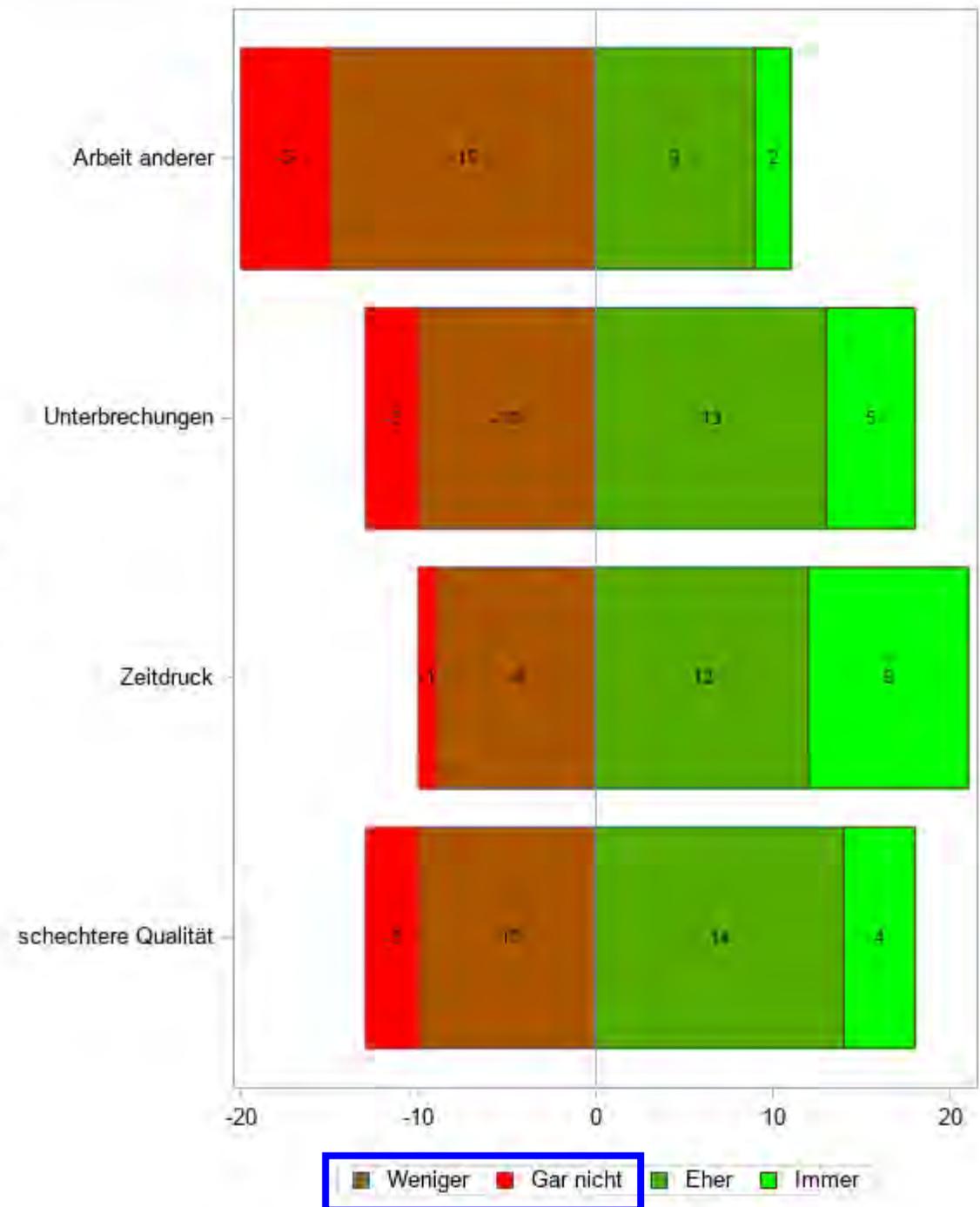
Ursache



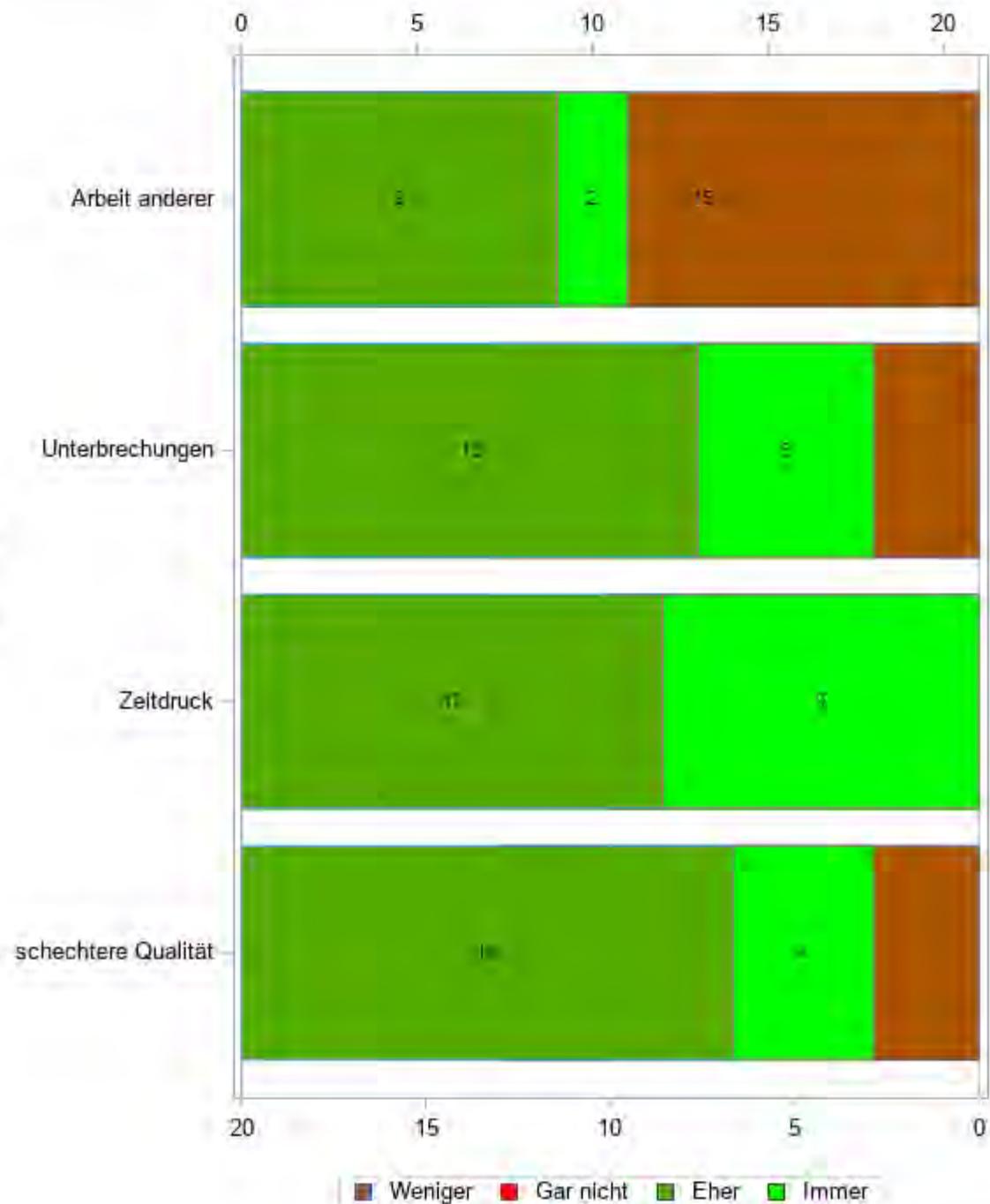
Ursache



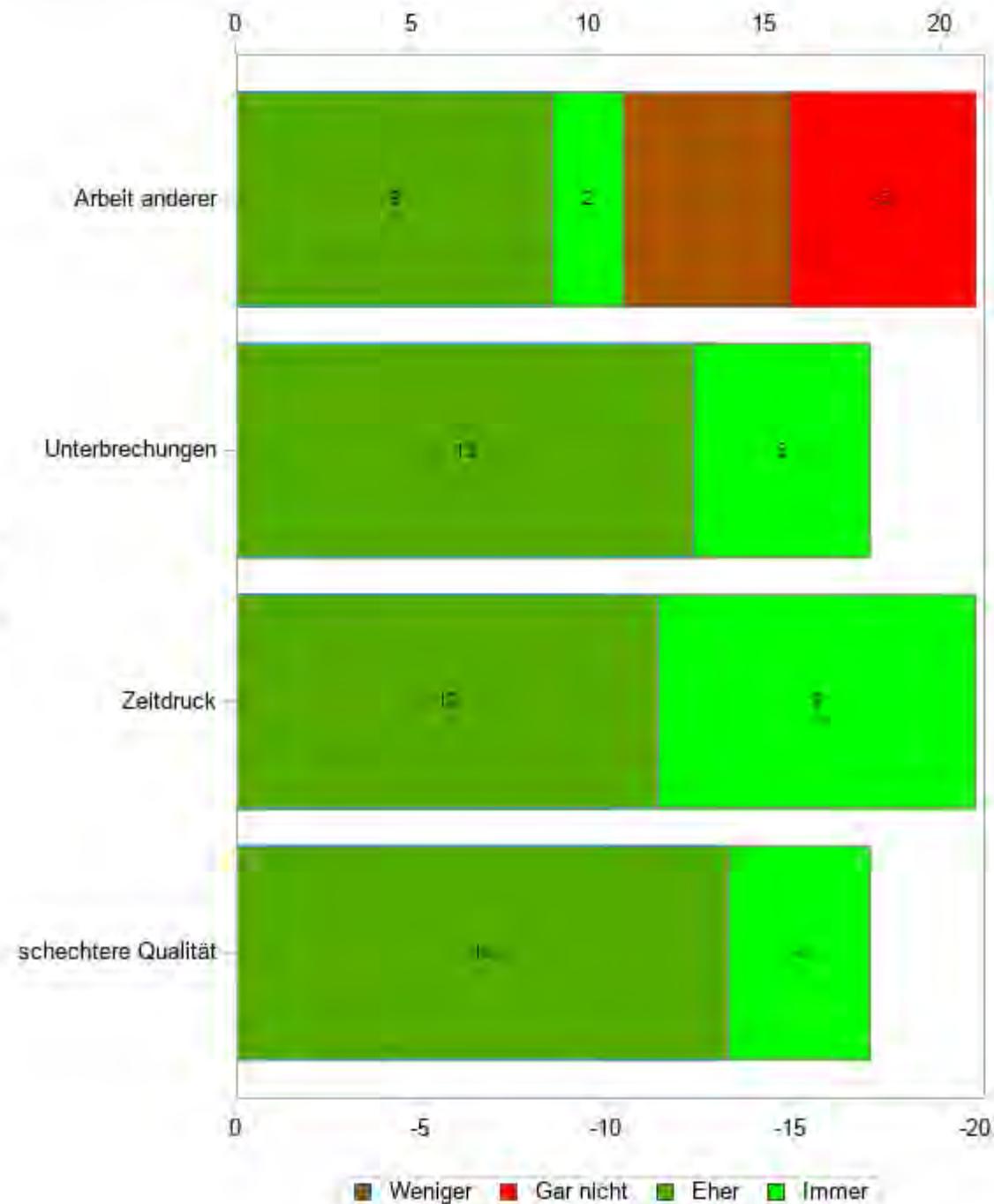
Ursache



Ursache



Ursache



sgplot



```
title Height=&tsize.pt &title.;

proc sgplot data=&daten.2 noautolegend dattrmap=&range.3 &sganno.; /* Zuweisung der Farbtabelle, keine automatische Legende,
                                                                    goupdisplay=stack ist default */
hbar &source. / name='A' /* Horizontale Bar der linken Werte über die "&source.",
                                                                    Name für die Legende */

    response=Wert1 /* Welche Daten, da numerisch ist stat=sum default */
    group=&group. /* Gruppierung nach der Ausprägung */
    attrid=&group. /* Zuordnung &group. --> Farbe */
    grouporder=reversedata /* Sortierung der Segmente --> probieren, was passt */
    fill missing /* NOTWENDIG !!!, sonst funktioniert die Farbzuoordnung nicht */
    displaybaseline=&baseline.; /* Null-Linie wird gezeichnet oder nicht */
hbar &source. / x2axis name='B' /* Horizontale Bar der rechten Werte über die "&source.",
                                                                    Name für die Legende */

    response=Wert2 /* Welche Daten, da numerisch ist stat=sum default */
    group=&group. /* Gruppierung nach der Ausprägung */
    attrid=&group. /* Zuordnung &group. --> Farbe */
    grouporder=data /* Sortierung der Segmente --> probieren, was passt */
    fill missing /* NOTWENDIG !!!, sonst funktioniert die Farbzuoordnung nicht */
    displaybaseline=&baseline.; /* Null-Linie wird gezeichnet oder nicht */
axis display=(nolabel novalues) min=&min1. reverse max=&max1.; /* keine x1-Achsen-Werte, Orientierung rechts ->links,
                                                                    Minimum ist das Maximum der rechten Summe */
x2axis display=none min=&min2. max=&max2.; /* keine x2-Achse, Minimum ist das Maximum der linken Summe */
yaxis display=(nolabel) values=(&yvalues.) fitpolicy=&split. valueattrs=(size=&textsize.);
                                                                    /* Keine Beschriftung der y-Achse , Beachtung der Reihenfolge */
keylegend 'B' / exclude=(&leer_wert.) valueattrs=(size=&textsize.); /* Legende anhand der Sortierung der rechten Werte */
run;
```

```
proc sgplot data=&daten.2 noautolegend datrmap=&range.3 &sganno. ;
hbar &source. / name='A'
    response=Wert1
    group=&group.
    attrid=&group.
    grouporder=reversedata
    fill missing
    displaybaseline=&baseline. ;
hbar &source. / x2axis name='B'
    response=Wert2
    group=&group.
    attrid=&group.
    grouporder=data
    fill missing
    displaybaseline=&baseline. ;
xaxis display=(nolabel novalues) min=&min1. reverse max=&max1. ;
x2axis display=none min=&min2. max=&max2. ;
yaxis display=(nolabel) values=(&yvalues.) fitpolicy=&split.
valueattrs=(size=&textsize. ) ;
keylegend 'B' / exclude=(&leer_wert.) valueattrs=(size=&textsize. ) ;
run ;
```

Herausforderungen



1. Programmsteuerung --> Makrovariablen inkl. `sysevalf`, `sysfunc`
2. Aufteilung links, mitte, rechts --> `&source.`
3. Farbzuzuweisung --> `&range.3`
4. Richtiger Nullpunkt --> `&baseline.`
5. einheitliche Skalierung links <-> rechts --> `&min1.` `&max1.` `&min2.` `&max2.`
6. Sortierung der Daten --> `&daten.2`
7. Seglabel, speziell Mitte --> `&sganno.` `&leer_wert.`
8. Legende --> `exclude`
9. ODS Graphics / `width<>640` --> `moving`

1. Programmsteuerung



```
* ----- ;
* Steuerungsparameter ;
* ----- ;

/* Korrektur der Positionen der Segmentlabel, notwendig,
   wenn Standardgrafikbreite (640) geändert wird */
%LET height=480;          /* Standard=480 */
%LET width=640;          /* Standard=640 */
%LET moving=0;           /* Korrektur der Segmentlabel-Position, 640-->0,
                           alles andere mit Probieren */

%LET textsize = 10;      /* Größe der Texte der Legende und Y-Achse */
%LET split = split;     /* Splitting der Y-Achsen-Werte, mögliche Werte:
                           NONE | SPLIT | SPLITALWAYS | THIN */

/* Titel und Größe in der Grafik */
%LET Title = ;
%LET TSize = 12;
```

1. Programmsteuerung



```
/* Rechnen mit den Werten (Prozente=Off)
   oder den Prozentwerten (Prozente=On) */
%LET Prozente = Off;

/* Automatische Farbzweisung und Sortierung (AutoCode=On)
   oder im Excel-Sheet (AutoCode=Off) */
%LET AutoCode = On;

/* Name=<file> und Datenblatt=<sheet1> der Exceldatei */
%LET file="D:\Biometrie\Kurse\2022\2022 - Fragebogen\Likert\Likert7.xlsx";
%LET sheet1="Daten";
```

1. Programmsteuerung

```
/* Wenn AutoCode=Off dann hier den Namen des Sheets  
mit den Rängen, Faktoren und Farben */
```

```
/* ID, VALUE und FILLCOLOR für die Farben, Namen müssen so bleiben */
```

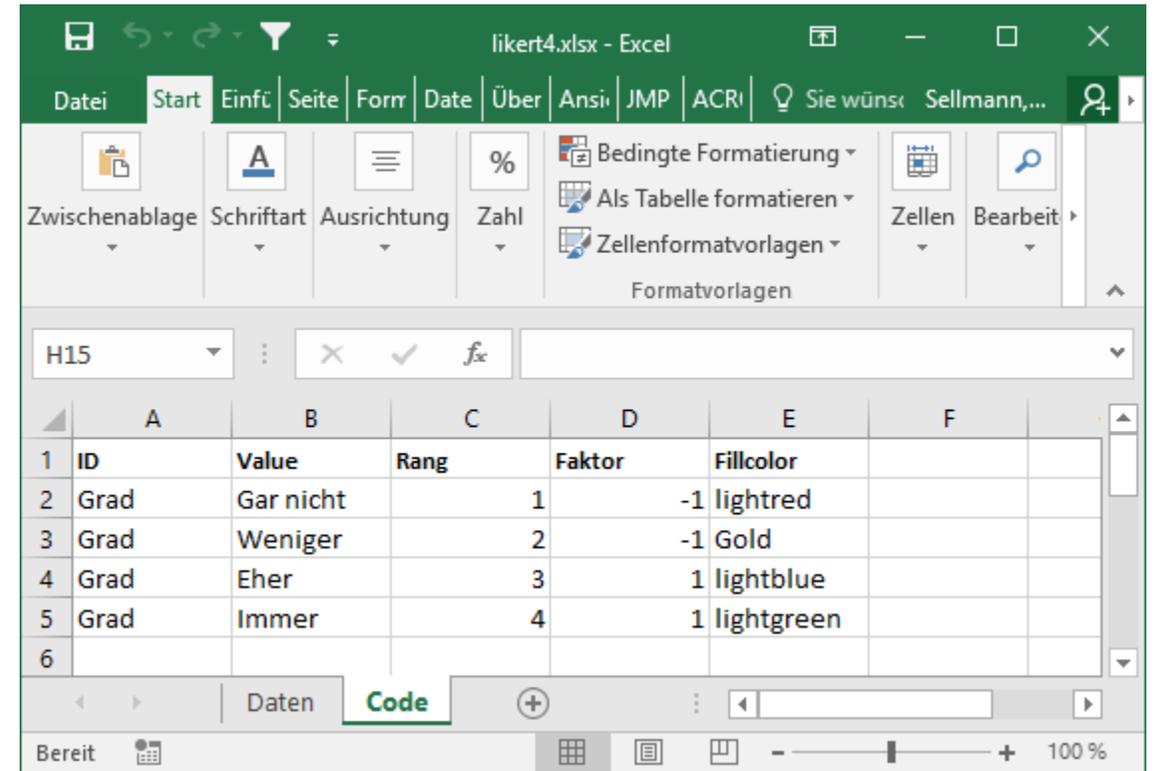
```
/* Wert für ID darf nur einen Wert annehmen */
```

```
/* Werte in VALUE müssen mit den Werten der 1. Zeile  
(Spalte 2..N) des Daten-Sheets  
übereinstimmen */
```

```
/* FAKTOR für die Aufteilung in links  
(-1), Mitte bei ungeraden (0) und  
rechts (1), Name muss bleiben */
```

```
/* RANG für die Sortierung von links  
nach rechts, Name muss bleiben */
```

```
%LET sheet2="Code";
```

A screenshot of the Microsoft Excel interface. The title bar shows 'likert4.xlsx - Excel'. The ribbon is set to 'Start'. The active cell is H15. The data table is as follows:

	A	B	C	D	E	F
1	ID	Value	Rang	Faktor	Fillcolor	
2	Grad	Gar nicht		1	-1 lightred	
3	Grad	Weniger		2	-1 Gold	
4	Grad	Eher		3	1 lightblue	
5	Grad	Immer		4	1 lightgreen	
6						

1. Programmsteuerung

```
/* Größe und Farbe für die Segmente,  
   seglabel=On --> Anzeige der Werte in den Segmenten */  
%LET seglabel = On;  
%LET zahl_farbe = Black;  
%LET zahl_size   = 10;  
  
/* showNull=On --> Nullwerte anzeigen,  
   Off --> Nullwerte nicht anzeigen */  
%LET showNull = Off;  
  
/* Baseline zeichnen (On) oder nicht (Off) */  
%LET baseline = On;  
  
/* unique=On --> Null genau in der Mitte,  
   Off --> Null nach links bzw. rechts verschieben */  
%LET unique=Off;
```

1. Programmsteuerung



```
*-----;
* Weitere Parameter, am besten so stehen lassen ;
*-----;
/* Start, Mitte und Ende für Farbskala */
%LET start=red;
%LET mid=lightyellow;
%LET end=darkgreen;

/* Wert und Farbe für die Auffüll-Kategorie */
%LET leer_wert = " ";
%LET leer_farbe = "white";

/* Namen und Spalten der Ziel-Datasets */
%LET filename=%sysfunc(scan(&file., -1, "\"));
%LET daten=%sysfunc(scan(&filename., 1, "."));
%LET range=Range;
%LET group=Grad;
```

1. Programmsteuerung

`/* Einlesen Excel (dbms=xlsx), ohne Assistent, dieses geht immer */`

```
proc import out=&daten.
```

```
  file=&file.
```

```
  dbms=xlsx replace;
```

```
  sheet=&sheet1.;
```

```
run;
```

`/* Name der ersten Spalte */`

```
proc sql noprint;
```

```
select COALESCE(name, 'unknown') into :source
```

```
from dictionary.columns
```

```
where libname = 'WORK'
```

```
  and memname = upper("&daten.")
```

```
  and varnum = 1;
```

```
quit;
```

	A	B	
1		Komplett unzufrieden	Me
2	Frage 1		0
3	Frage 2		4
4	Frage 3		7
5	Frage 4		3
5	Frage 5		8
7	Frage 6		2
3	Frage 7		1
3	Frage 8		5

2. Aufteilung links, mitte, rechts

```
/* Label, Rang und Faktor der Spalten 2..N */  
proc sql noprint;  
select max(varnum) into :max from dictionary.columns  
where libname = 'WORK' and memname = upper("&daten.")  
and varnum>1 and label is not null;  
quit;
```

```
proc sql noprint;  
create table &range. as  
select "&group." as ID  
label as Value  
varnum-1 as Rang  
(2*(varnum-1)) as Faktor  
from dictionary.columns  
where libname = 'WORK' and  
and varnum>1 and label  
order by varnum;  
quit;
```

	ID	Value	Rang	Faktor
1	Grad	Komplett unzufrieden	1	-1
2	Grad	Meist unzufrieden	2	-1
3	Grad	Etwas unzufrieden	3	-1
4	Grad	weder zufrieden noch unzufrieden	4	0
5	Grad	Etwas zufrieden	5	1
6	Grad	Meist zufrieden	6	1
7	Grad	Rundum zufrieden	7	1

3. Farbzuzuweisung

```

/* Colorramp --> generiert die Farben */
/* http://euclid.psych.yorku.ca/datavis/sasmac/colorramp.html */
%include "D:\Biometrie\Kurse\2022\2022 - Fragebogen\Likert\Colorramp.sas";

/* Wieviele Ränge */
%LET raenge=%sysevalf(&max.-1);

/* N=Raenge Farben auswählen --> work.Colors */
/* BUG (?) in colorramp bei ungeradem Range wird immer 2 addiert */
%LET bug=%sysevalf(2-(&raenge%2));
%colorramp(start=&start, n=&raenge, factor=&factor, fillcolor=&fillcolor);

/* Farben entsprechen Rang */
proc sql noprint;
create table &range.
select a.*, b.color
from &range. a, color b
where a.rang=b.colorrank+&bug;
quit;

```

VIEWTABLE: Work.Range1

	ID	Value	Rang	Faktor	FillColor
1	Grad	Komplett unzufrieden	1	-1	CXFF0000
2	Grad	Meist unzufrieden	2	-1	CXFF551D
3	Grad	Etwas unzufrieden	3	-1	CXFFAA39
4	Grad	weder zufrieden noch unzufrieden	4	0	CXFFFF56
5	Grad	Etwas zufrieden	5	1	CXAAE839
6	Grad	Meist zufrieden	6	1	CX55D21D
7	Grad	Rundum zufrieden	7	1	CX00BB00

3. Farbzuzuweisung



```
%macro AutoCode;
```

```
/* Alternativ: Farben und Ränge im separaten Sheet festlegen */
```

```
/* Für die Reihenfolge und die Farben in der Grafik */
```

```
%if %sysevalf(%sysfunc(upcase("&AutoCode.")) eq "OFF")
```

```
%then %do;
```

```
proc import out=&range.1 file=&file.
```

```
dbms=xlsx replace;
```

```
sheet=&sheet2.;
```

```
run;
```

```
/* Eindeutigen Wert für ID bestimmen */
```

```
proc sql noprint; select max(id) into :group
```

```
from &range.1; quit;
```

```
%end;
```

```
%mend;
```

```
%AutoCode;
```

3. Farbzuzuweisung



```
/* Farbe für die Leerwerte */  
data &range.2;  
retain ID "&group." FillColor &leer_farbe.;  
Value=&leer_wert.; output;  
run;
```

	ID	FillColor	Value
1	Grad	white	

```
/* Auffüllfarbe hinzufügen */  
data &range.3;  
set &range.1 &range.2  
run;
```

	ID	Value	Rang	Faktor	FillColor
1	Grad	Komplett unzufrieden	1	-1	CXFF0000
2	Grad	Meist unzufrieden	2	-1	CXFF551D
3	Grad	Etwas unzufrieden	3	-1	CXFFAA39
4	Grad	weder zufrieden noch unzufrieden	4	0	CXFFFF56
5	Grad	Etwas zufrieden	5	1	CXAAE839
6	Grad	Meist zufrieden	6	1	CX55D21D
7	Grad	Rundum zufrieden	7	1	CX00BB00
8	Grad		.	.	white

4. **Richtiger Nullpunkt**
5. **einheitliche Skalierung links und rechts**
6. **Sortierung der Daten**
7. **Seglabel, speziell Mitte**

Bisher noch sehr übersichtlich und strukturierbar.

Die Punkte 4. bis 7. lassen sich jetzt aber nicht mehr so klar trennen,
Deshalb werden sie jetzt in einem Block „4.-7. Datentransformation“ behandelt.

Außerdem wird es jetzt kompliziert.

4.-7. Datentransformation



```
/* Merken der Original-Reihenfolge */  
  
/* Daten aus Tabellenform in SAS-Form transponieren */  
  
/* Prozente statt Werte? */  
%macro Prozente; %mend; %Prozente;  
  
/* Aufteilung der Werte in 2 Spalten,  
links von der Null -> Wert1,  
rechts von der Null -> Wert2 */  
  
/* Faktor=0 --> halbieren nach links und rechts */  
  
/* negatives Maximum der Summen der linken und rechten Seite für die  
Nullpunktbestimmung, speichern in Makrovariablen min1 und min2 */
```

4.-7. Datentransformation



```
/* Den kleineren merken --> min1 in xaxis, min2 in x2axis */  
  
/* Werte für die y-Achse in Originalreihenfolge */  
  
/* links kleiner */  
  
/* Rechts kleiner */  
  
/* Die "kleineren" auffüllen, es reicht ein (größter) Wert */  
  
/* Daten inkl. Auffüllwerten in richtiger Reihenfolge  
sortieren --> grouporder */
```

4.-7. Datentransformation



```
/* Die Positionen im Annotate-Datenatz festlegen */

/* Position ist der Mittelwert von linker und rechter Begrenzung */
data anno2;
retain function "text" drawspace "datavalue"
    anchor "center" textcolor "&zahl_farbe."
    textsize &zahl_size.;
set temp_wert21;
rename &source.=ycl1;
length label $12;
if faktor=0
    then do; x1=0+&moving.; label=2*wert2; end;
    else do; x1=((summe-wert2)+summe)/2-&moving.; label=wert2; end;
run;

/* Der gemeinsame Datensatz, Faktor=0 nur aus der rechten Seite, die
    rechten Positionen jetzt negativ */
```

4.-7. Datentransformation



```
/* Seglabel anzeigen */  
%macro seglabel;  
    %global sganno;  
  
    %if %sysevalf(%sysfunc(upcase("&seglabel.")) eq "ON")  
        %then %LET sganno=sganno=anno;  
        %else %LET sganno=;  
  
%mend;  
  
%seglabel;
```

4.-7. Datentransformation

```
%put &daten.2;
%put &source.;
```

Likert72
A



VIEWTABLE: Work.Likert72

	A	Nr	Grad	Rang	Faktor	Wert1	Wert2
1	Frage 7	0		0	.	81	.
2	Frage 1	1	Komplett unzufrieden	1	-1	0	.
3	Frage 2	2	Komplett unzufrieden	1	-1	4	.
4	Frage 3	3	Komplett unzufrieden	1	-1	7	.
24	Frage 7	7	Etwas unzufrieden	3	-1	8	.
25	Frage 8	8	Etwas unzufrieden	3	-1	17	.
26	Frage 1	1	weder zufrieden noch unzufrieden	4	0	11.5	11.5
27	Frage 2	2	weder zufrieden noch unzufrieden	4	0	15	15
28	Frage 3	3	weder zufrieden noch unzufrieden	4	0	13	13
29	Frage 4	4	weder zufrieden noch unzufrieden	4	0	10	10
30	Frage 5	5	weder zufrieden noch unzufrieden	4	0	15	15
31	Frage 6	6	weder zufrieden noch unzufrieden	4	0	22.5	22.5
32	Frage 7	7	weder zufrieden noch unzufrieden	4	0	20.5	20.5
33	Frage 8	8	weder zufrieden noch unzufrieden	4	0	13.5	13.5
34	Frage 1	1	Etwas zufrieden	5	1	.	50
35	Frage 2	2	Etwas zufrieden	5	1	.	61
36	Frage 3	3	Etwas zufrieden	5	1	.	37

4.-7. Datentransformation



```
%put &sganno;
```

sganno=anno

VIEWTABLE: Work.Anno

	function	drawspace	anchor	textcolor	textsize	yc1	Rang	Faktor	Wert1	summe	label	x1
1	text	datavalue	center	Black	10	Frage 1	3	-1	34	46	34	29
2	text	datavalue	center	Black	10	Frage 1	2	-1	8	54	8	50
3	text	datavalue	center	Black	10	Frage 1	1	-1	0	54	.	54
4	text	datavalue	center	Black	10	Frage 2	3	-1	28	43	28	29
5	text	datavalue	center	Black	10	Frage 2	2	-1	12	55	12	49
6	text	datavalue	center	Black	10	Frage 2	1	-1	4	59	4	57
7	text	datavalue	center	Black	10	Frage 3	3	-1	47	60	47	37
8	text	datavalue	center	Black	10	Frage 3	2	-1	10	70	10	65
9	text	datavalue	center	Black	10	Frage 3	1	-1	7	77	7	74
10	text	datavalue	center	Black	10	Frage 4	3	-1	33	43	33	27
11	text	datavalue	center	Black	10	Frage 4	2	-1	23	66	23	55
12	text	datavalue	center	Black	10	Frage 4	1	-1	3	69	3	68
13	text	datavalue	center	Black	10	Frage 5	3	-1	20	35	20	25

8. Legende

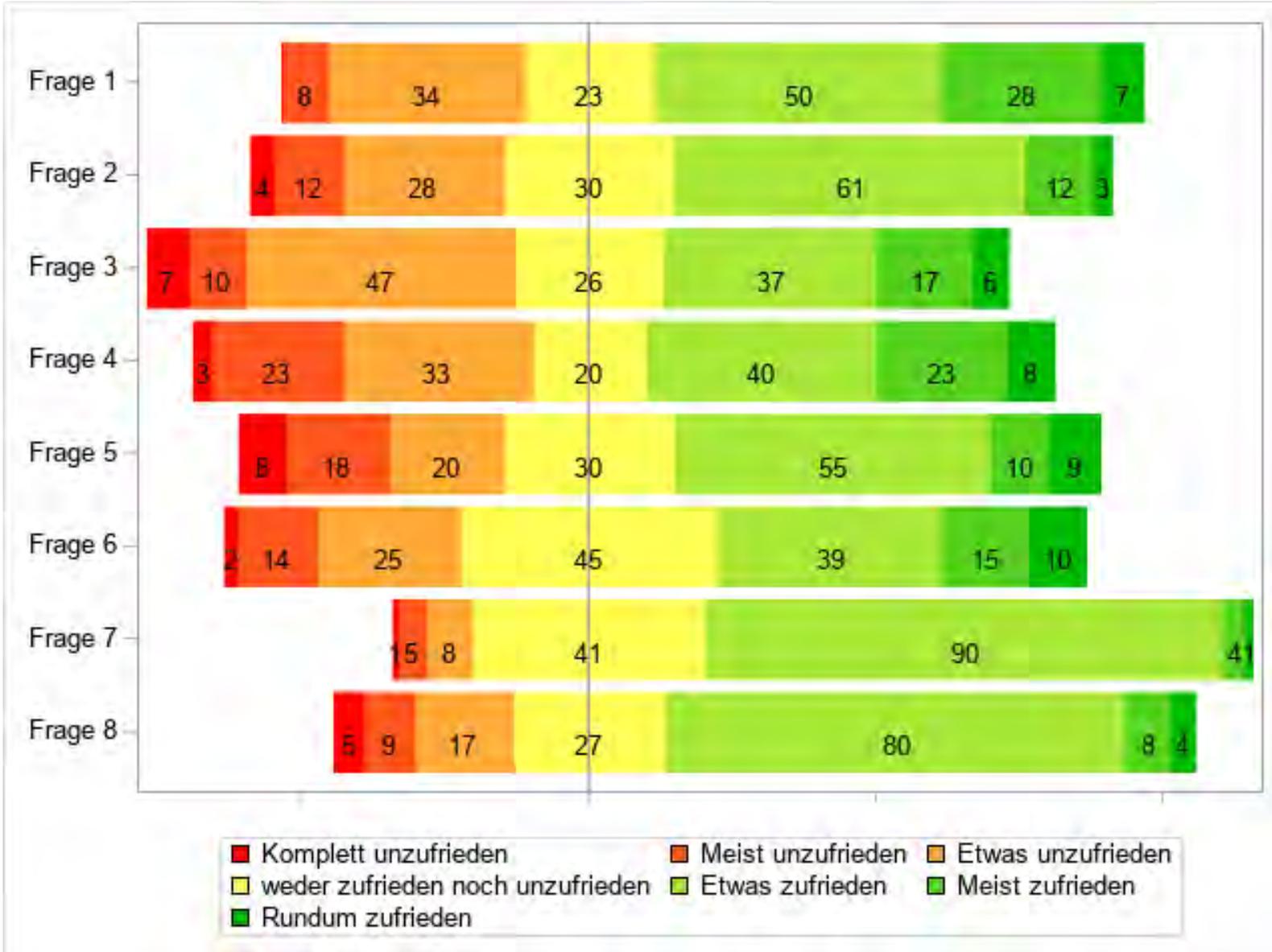


```
/* Legende anhand der Sortierung der rechten Werte */  
keylegend 'B' / exclude=(&leer_wert.) valueattrs=(size=&textsize.);
```

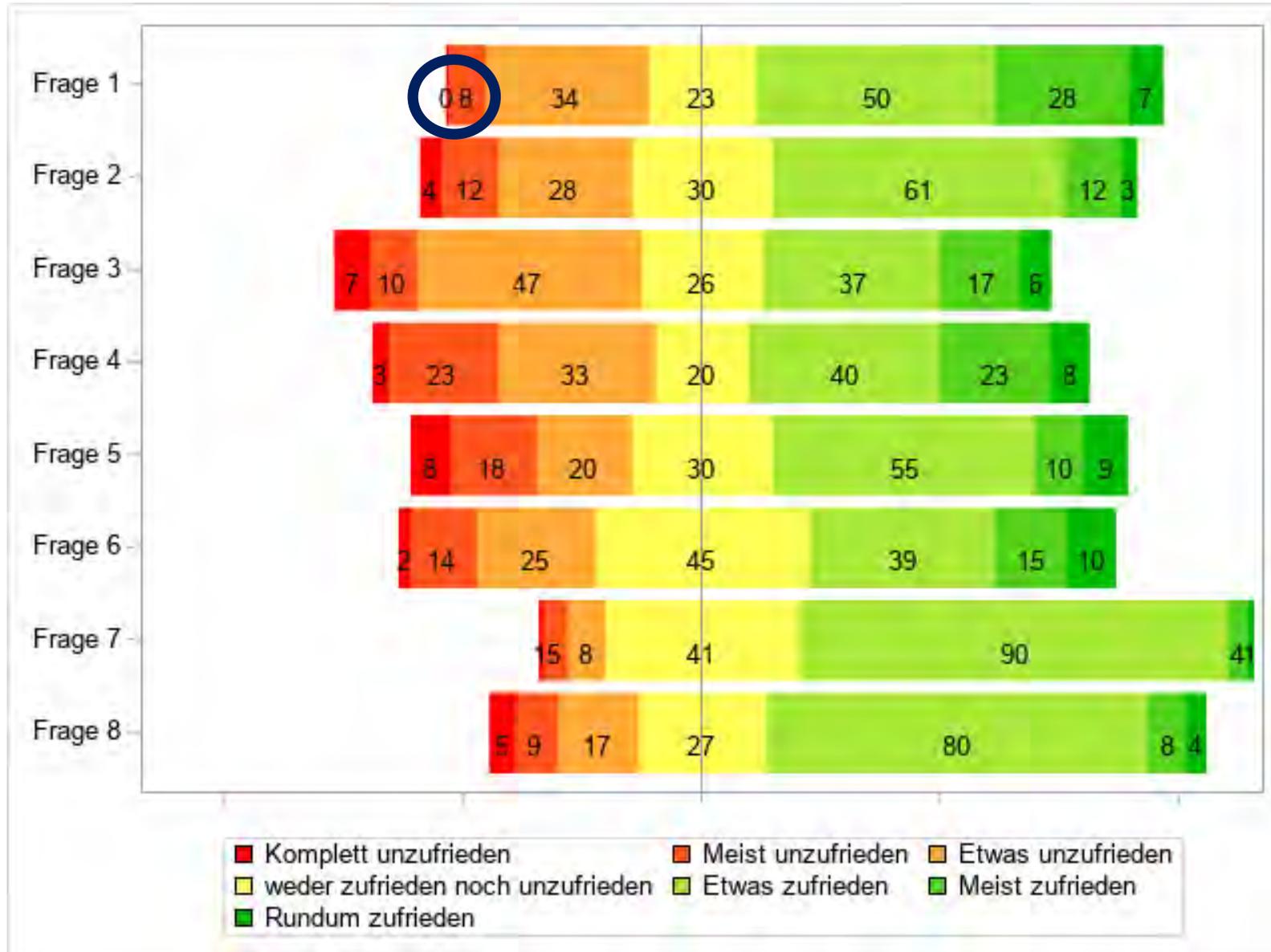
9. ODS Graphics / width<>640

Beispiel

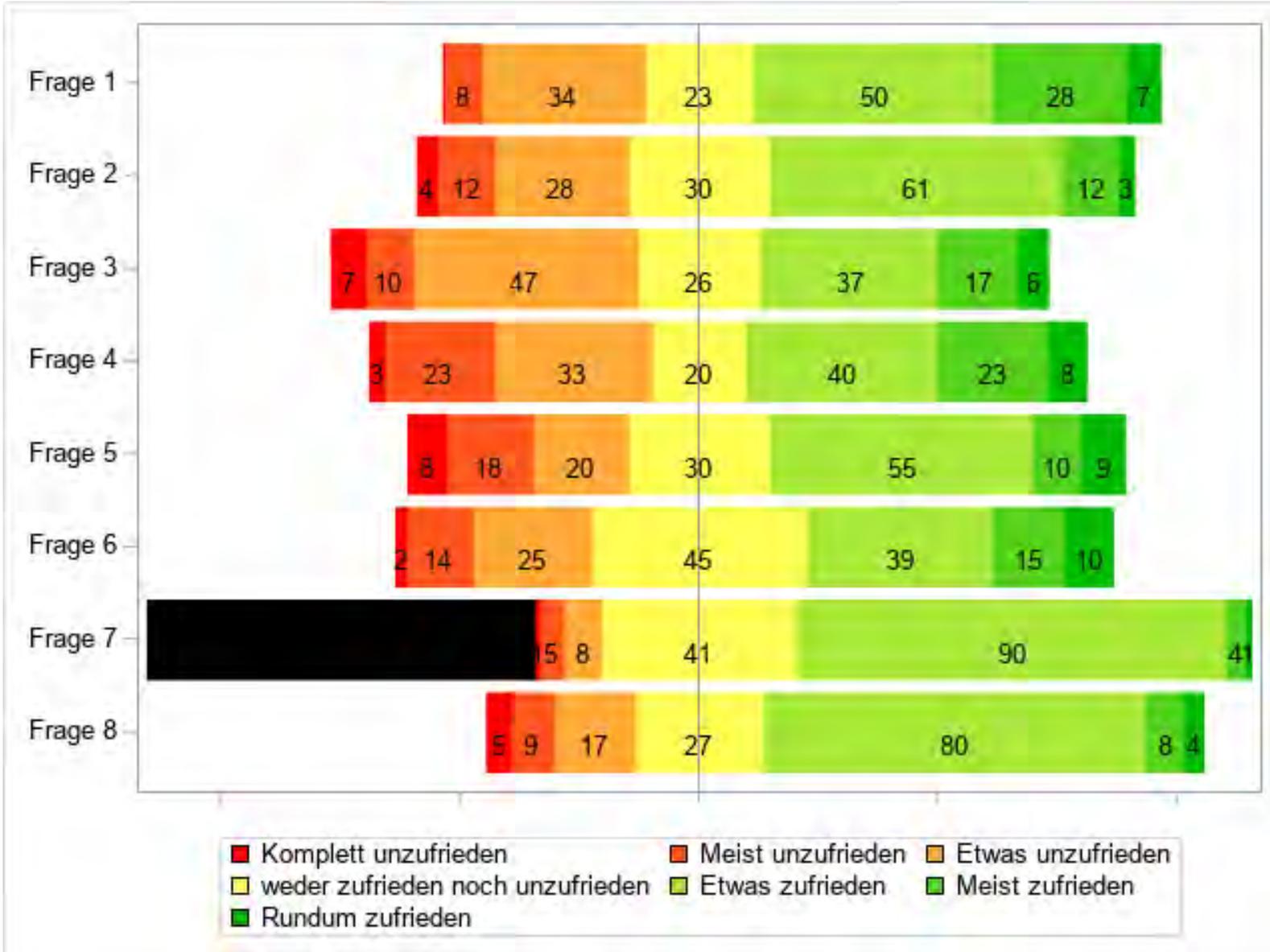
Bonus: Blick hinter die Kulissen



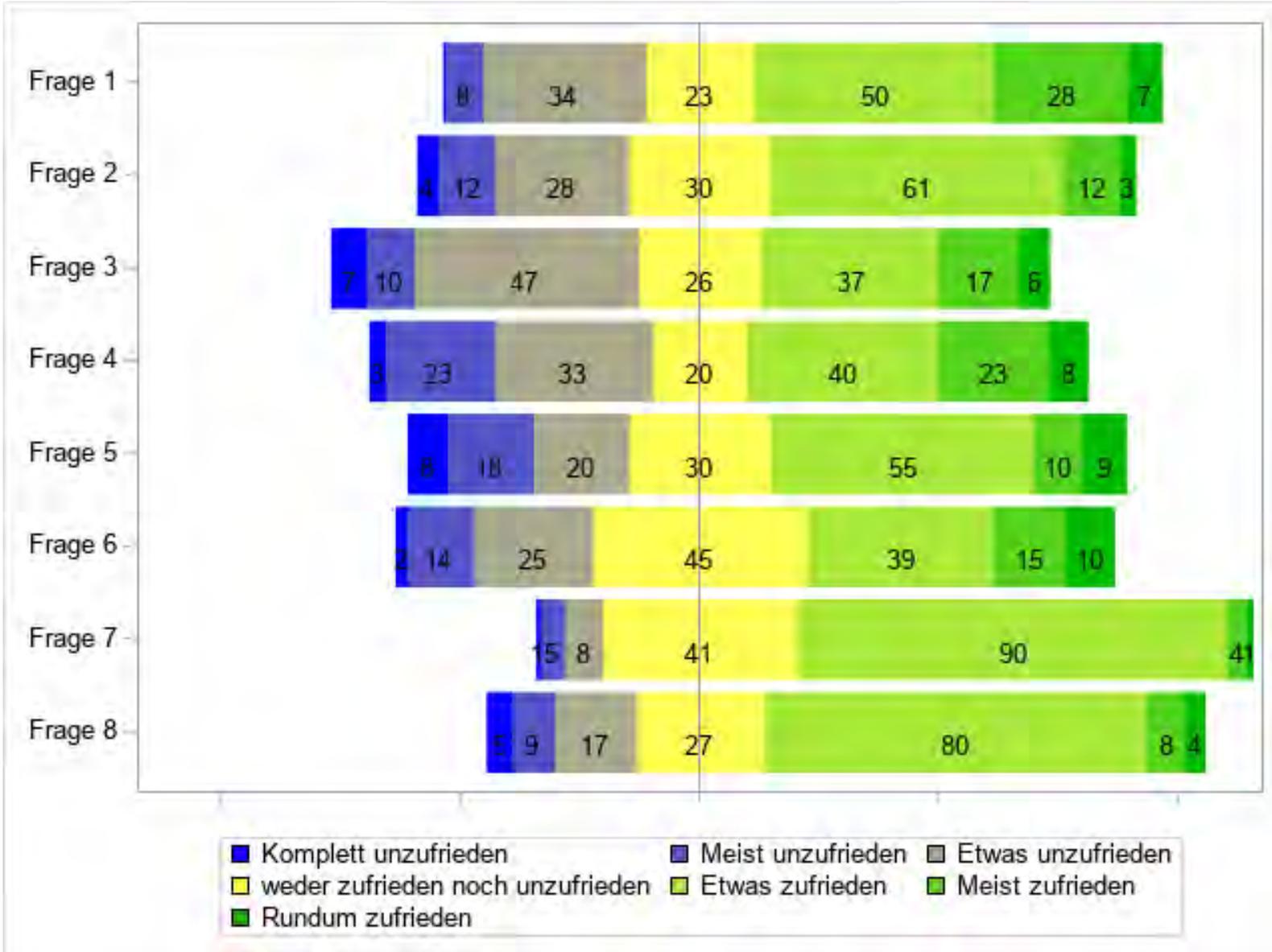
Bonus: Blick hinter die Kulissen



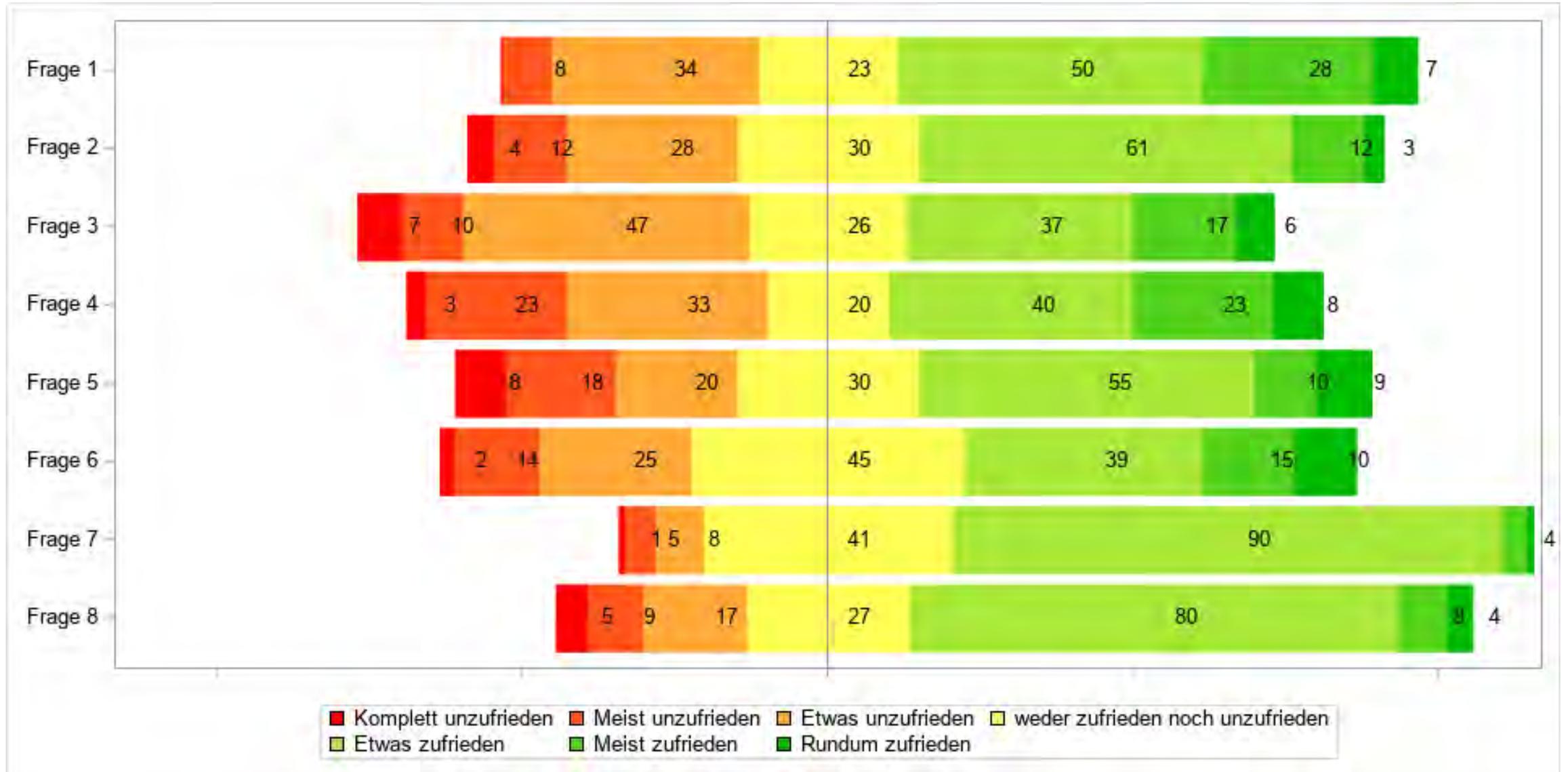
Bonus: Blick hinter die Kulissen



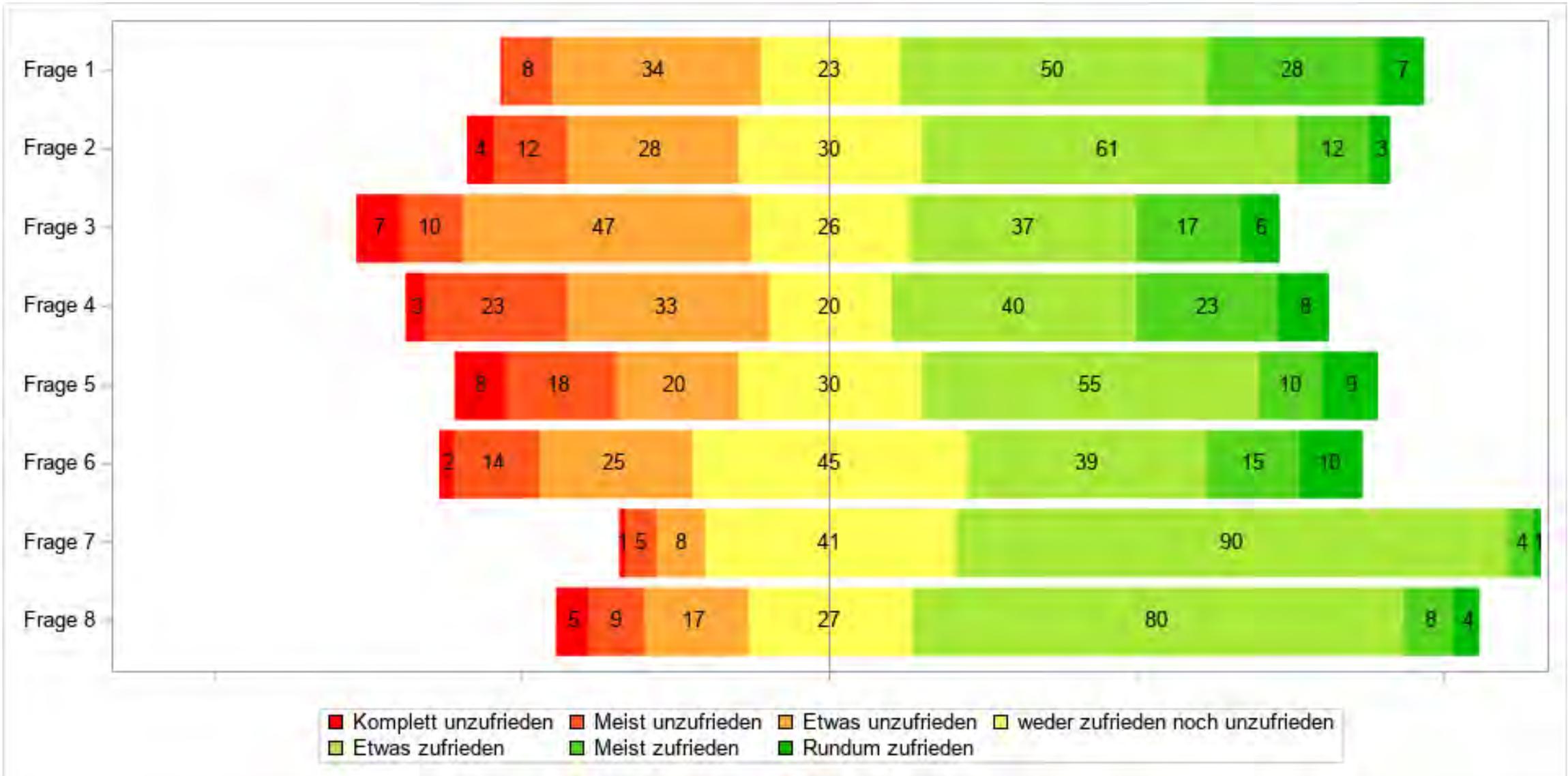
Bonus: Blick hinter die Kulissen



Bonus: Blick hinter die Kulissen



Bonus: Blick hinter die Kulissen



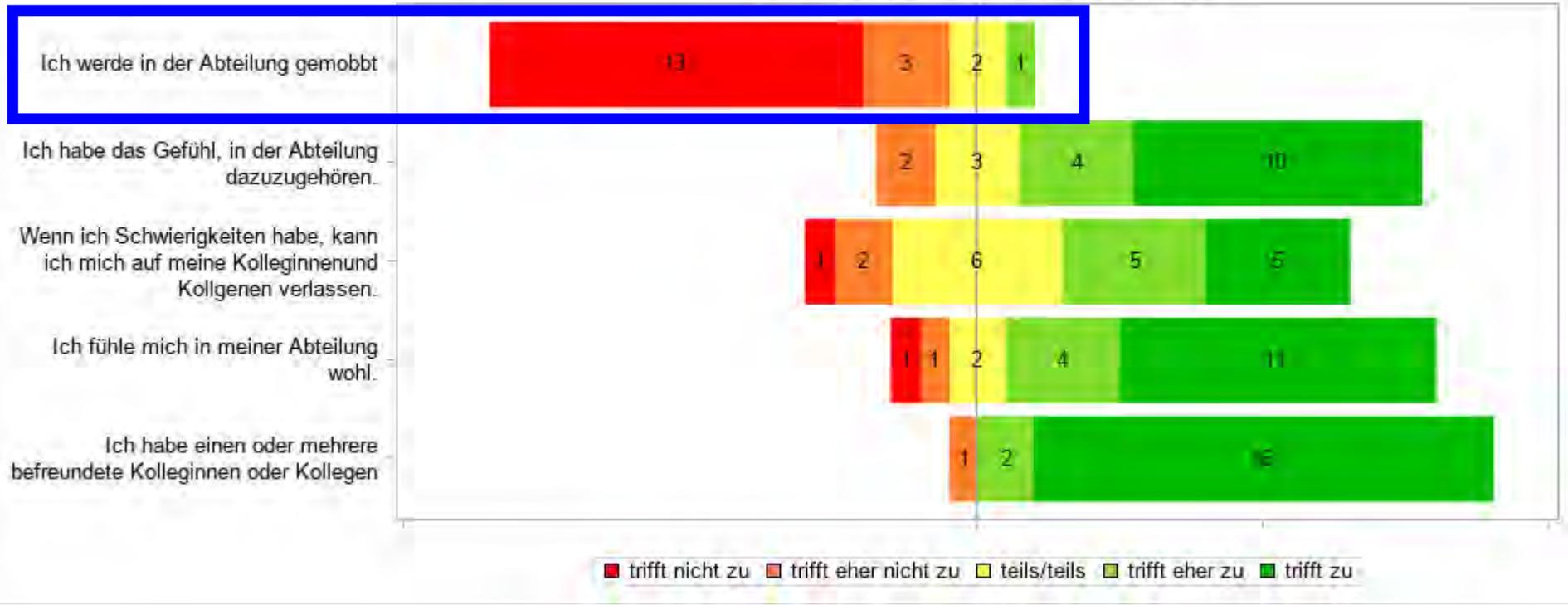
Die Noten wurden so vergeben (insgesamt 51 Antworten)



Ausblick



Wie geht es Dir in Deiner Abteilung?



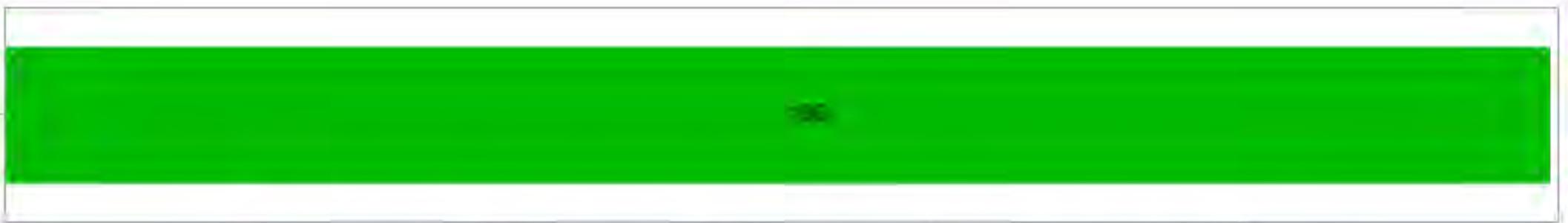
Problem: Was ist positiv und was ist negativ?

Ausblick



1. Richtige Fragen mit der richtigen Richtung
 - Ist aber (zum Glück) nicht das Problem von SAS
2. Höhe der Balken von SAS (oder von mir) nicht einstellbar
 - Wenn es mehrere Fragen-Blöcke mit unterschiedlicher Anzahl von Fragen gibt, aber die Höhe der SAS-Grafik gleich bleibt, dann sind die Balken innerhalb einer Grafik gleich hoch, zwischen den Grafiken aber unterschiedlich
 - Probieren mit ODS / height=
3. Formel für MOVING statt Probieren
 - Hängt von der Breite, aber offensichtlich auch von der Anzahl der Kategorien ab
4. Das gesamte Programm als ein Makro?
5. Download unter <http://sf.julius-kuehn.de/sas/2022/>

Wie
fanden
Sie
den
Vortrag?



- Kompletzt unzufrieden
- Meist unzufrieden
- Etwas unzufrieden
- weder zufrieden noch unzufrieden
- Etwas zufrieden
- Meist zufrieden
- Rundum zufrieden