

Update on Research into the use of  
Web Panel Survey Data by  
The Division of Research and  
Methodology (DRM) – **the RANDS  
Project**

Alan H. Dorfman, PhD

Paul Scanlon, PhD

Board of Scientific Counsellors May 19, 2016

# Outline

- Goals of the RANDS
- Data Collection
- Goal 1 Research to Date
  - Early Results
- Goal 2 Research to Date
  - The basic scheme/strategy
  - Main experiment planned
  - Preliminary experiment(s), in progress
- Questions

# Goals of the Research

- (1) To determine usefulness of web data in supplementing direct cognitive interviewing in determining how survey questions are interpreted by different groups
- (2) To assess the feasibility of replacing portions of traditional surveys with web surveys

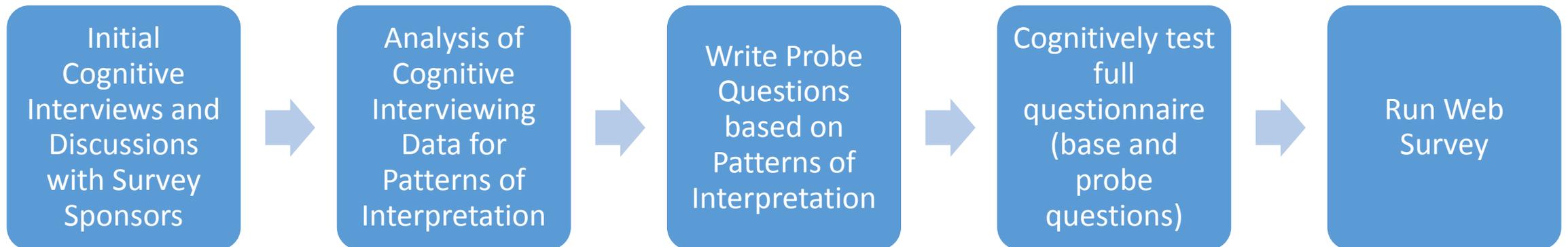
# Data Collection

- Contract with Gallup, currently for 2 Rounds of Web Panel Survey
  - Each round collects approximately 2,000 completed interviews
  - Panel is probability-based, drawn mainly from Gallup's RDD Daily Tracker
  - Sampling scheme specified by DRM
    - Strata include Race/Ethnicity, Age, Education
- Questionnaire is composed of NHIS questions (Round 1 and 2) and CQDER-designed cognitive probe question (Round 2 only)
- Gallup provides extra data on panel members, including demographics and weights

# **Goal One: Early Results using Targeted, Embedded Cognitive Probes**

## Targeted, Embedded Probes

- **Targeted: Not exploratory, content comes directly from cog. interviews**
- **Embedded: Placed in the survey instrument directly following Q of interest**



## **Utility of Targeted Embedded Probes**

- 1. Determining the extent of a problematic pattern of interpretation**
- 2. Examining patterns of interpretation across respondent groups**
- 3. Comparing the underlying constructs of two similarly-worded questions**

## Extent of a Problematic Pattern of Interpretation: Affect Question

- ANX\_1: How often [do/does] [you/he/she] feel worried, nervous or anxious? (Daily, Weekly, Monthly, A few times a year, Never)
  - ANX\_2: [Do/Does] [you/he/she] take medication for these feelings?
  - ANX\_3: Thinking about the last time [you/he/she] felt worried, nervous or anxious, how would [you/he/she] describe the level of these feelings? (A little, A lot, Somewhere in between a little and a lot)
- **Previous work on the Washington Group on Disability Statistics found that there are four major patterns of interpretation when respondents consider the following set of questions:**
    1. Intense feelings that cause physical distress
    2. Positive feelings that help productivity
    3. Negative feelings that interfere with life
    4. Diagnosed with anxiety by a medical professional

## Extent of a Problematic Pattern of Interpretation: Affect Question

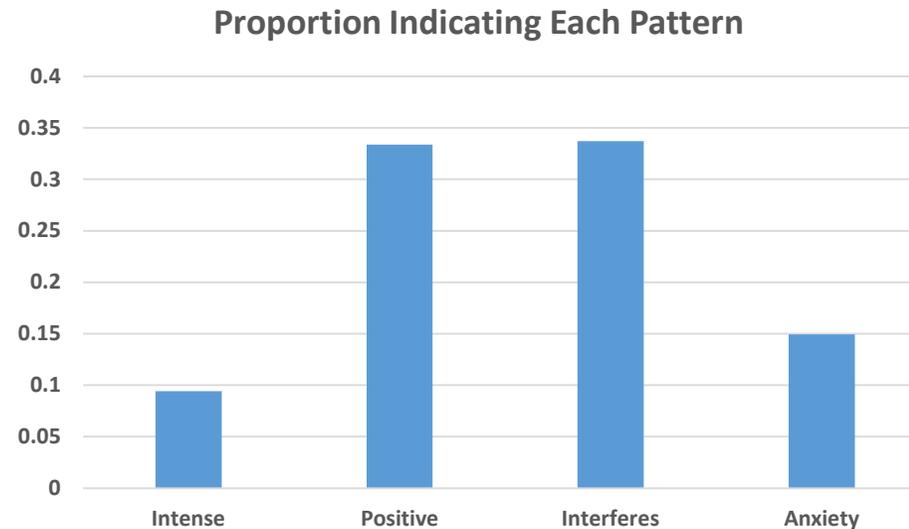
- ANX\_1: How often [do/does] [you/he/she] feel worried, nervous or anxious? (Daily, Weekly, Monthly, A few times a year, Never)
  - ANX\_2: [Do/Does] [you/he/she] take medication for these feelings?
  - ANX\_3: Thinking about the last time [you/he/she] felt worried, nervous or anxious, how would [you/he/she] describe the level of these feelings? (A little, A lot, Somewhere in between a little and a lot)
- **Previous work on the Washington Group on Disability Statistics found that there are four major patterns of interpretation when respondents consider the following set of questions:**
1. Intense feelings that cause physical distress
  2. Positive feelings that help productivity ← Out of Scope
  3. Negative feelings that interfere with life
  4. Diagnosed with anxiety by a medical professional

# Anxiety Probe

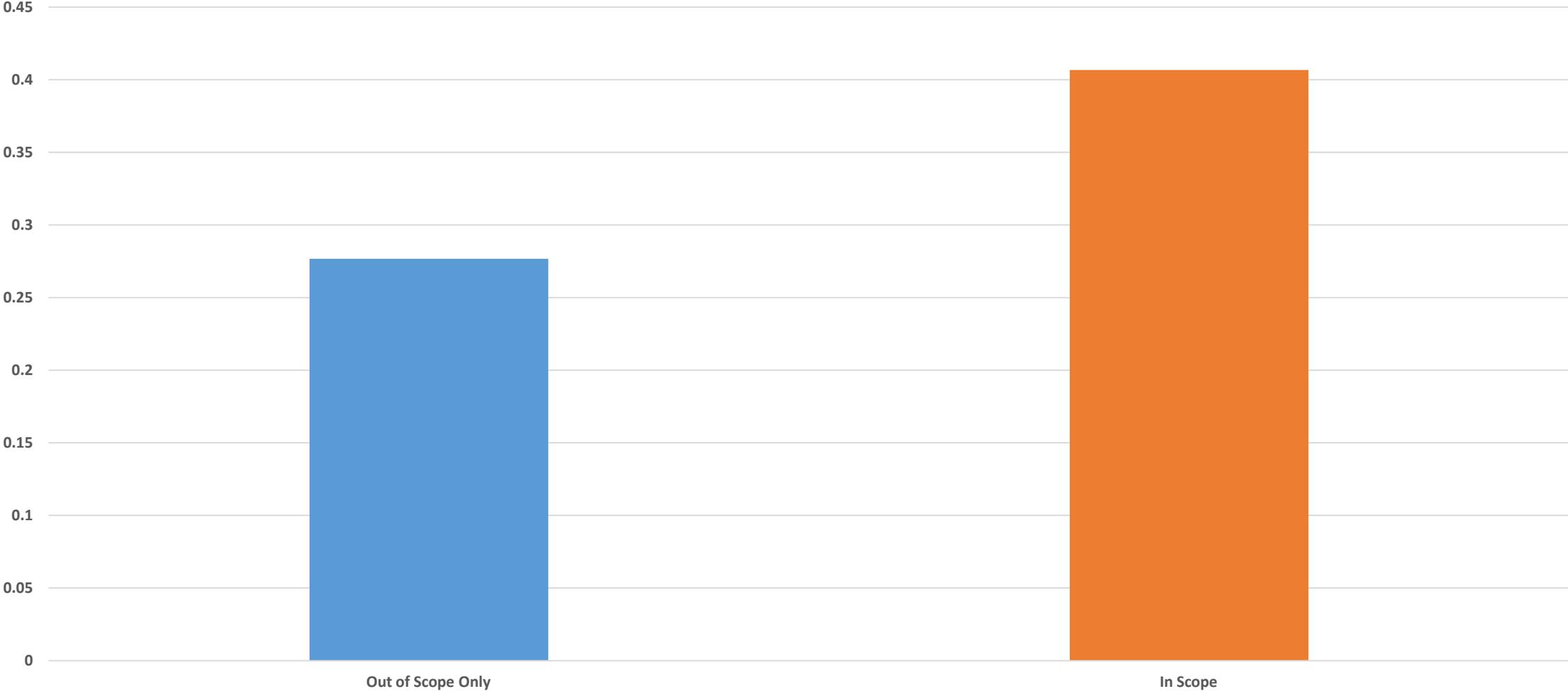
Which of the following statements, if any, describes your feelings? (Select all that apply)

1. Sometimes the feels can be so intense that my chest hurts and I have trouble breathing
2. These are positive feelings that help me to accomplish goals and be productive.
3. The feelings sometimes interfere with my life, and I wish that I did not have them.
4. I have been told by a medical professional that I have anxiety

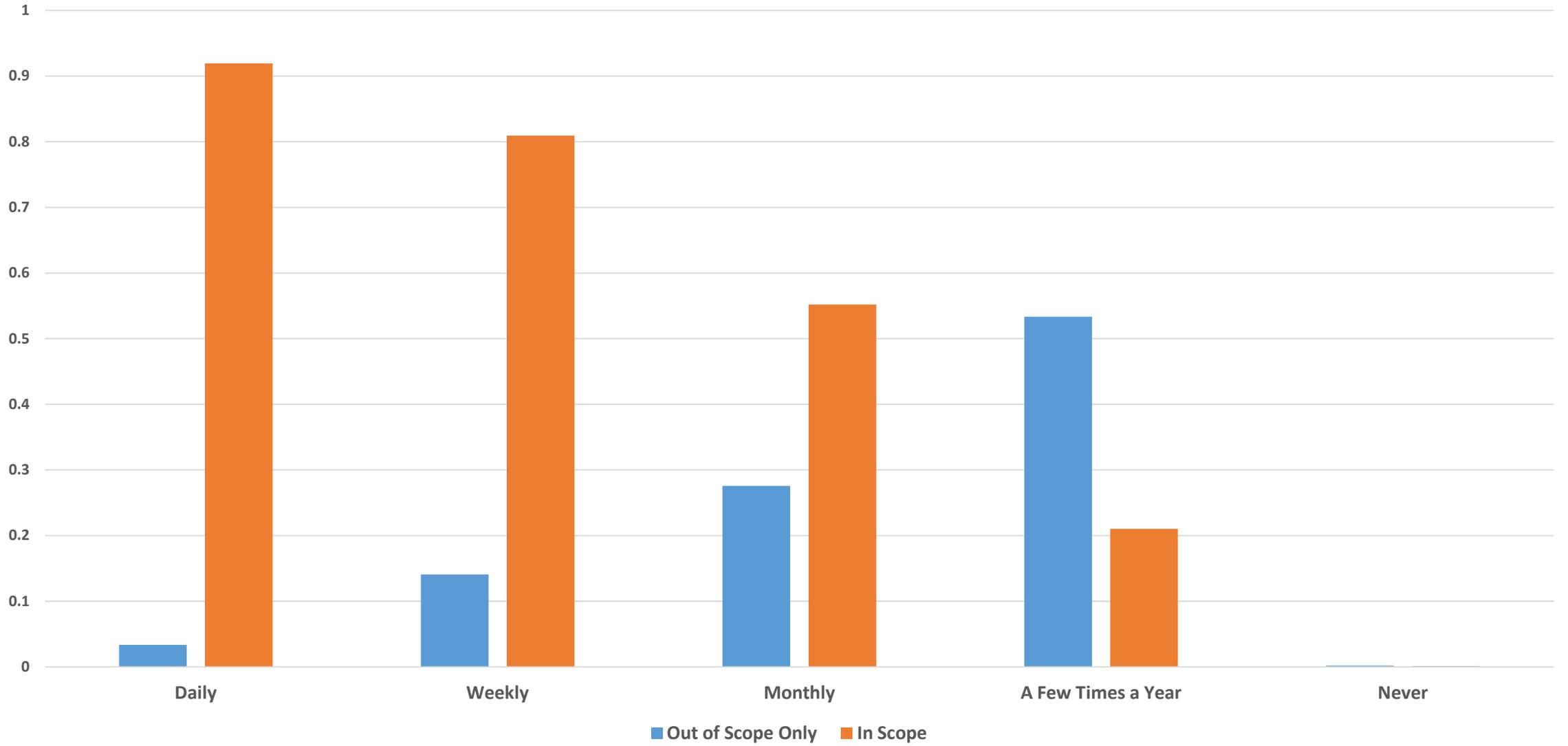
- **Basic distribution:**



Proportion Indicating an Out-of-Scope or In-Scope Pattern



Patterns by Response to Frequency Question



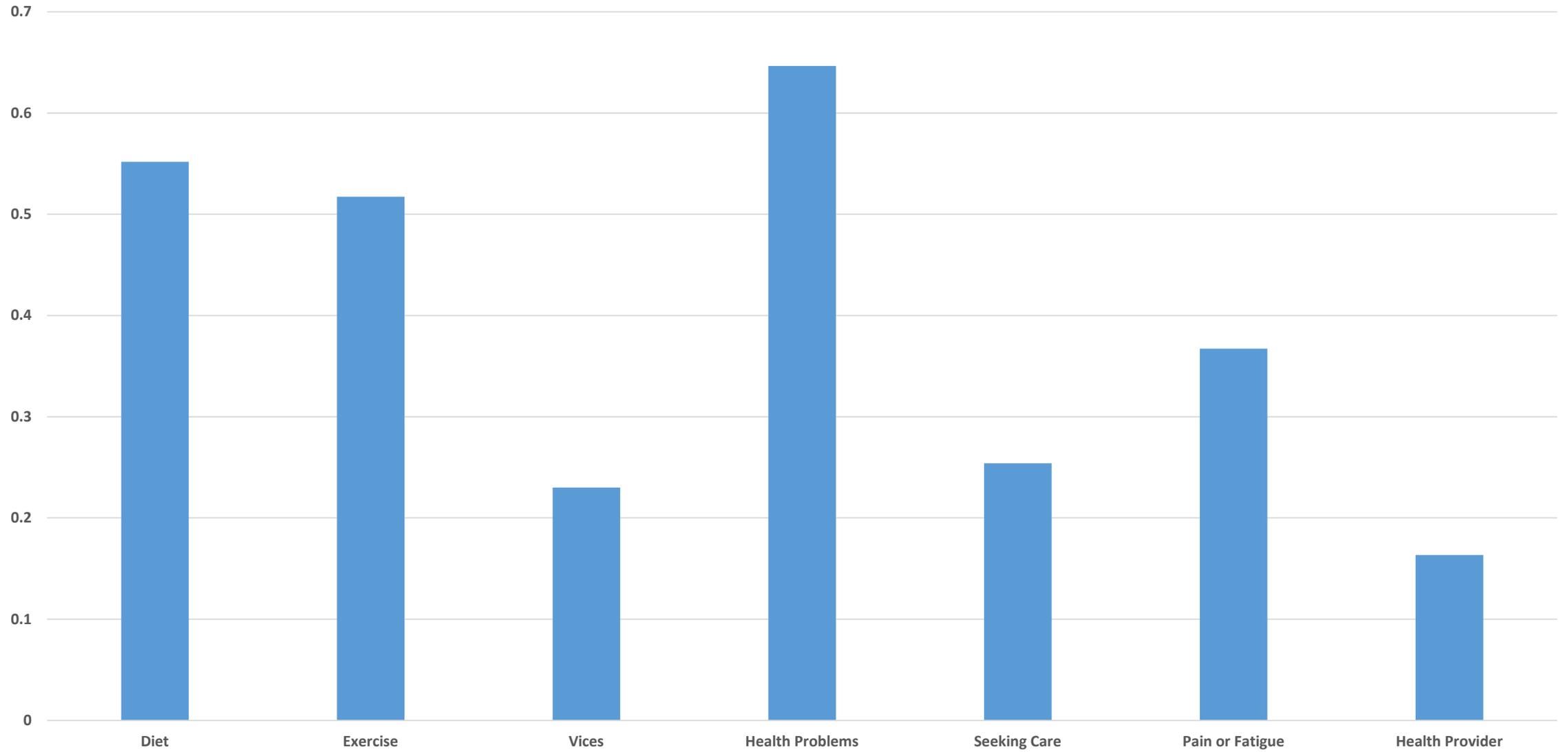
# Examining Patterns of Interpretation Across Respondent Groups: The General Health Status Question

Would you say your health in general is excellent, very good, good, fair, or poor?

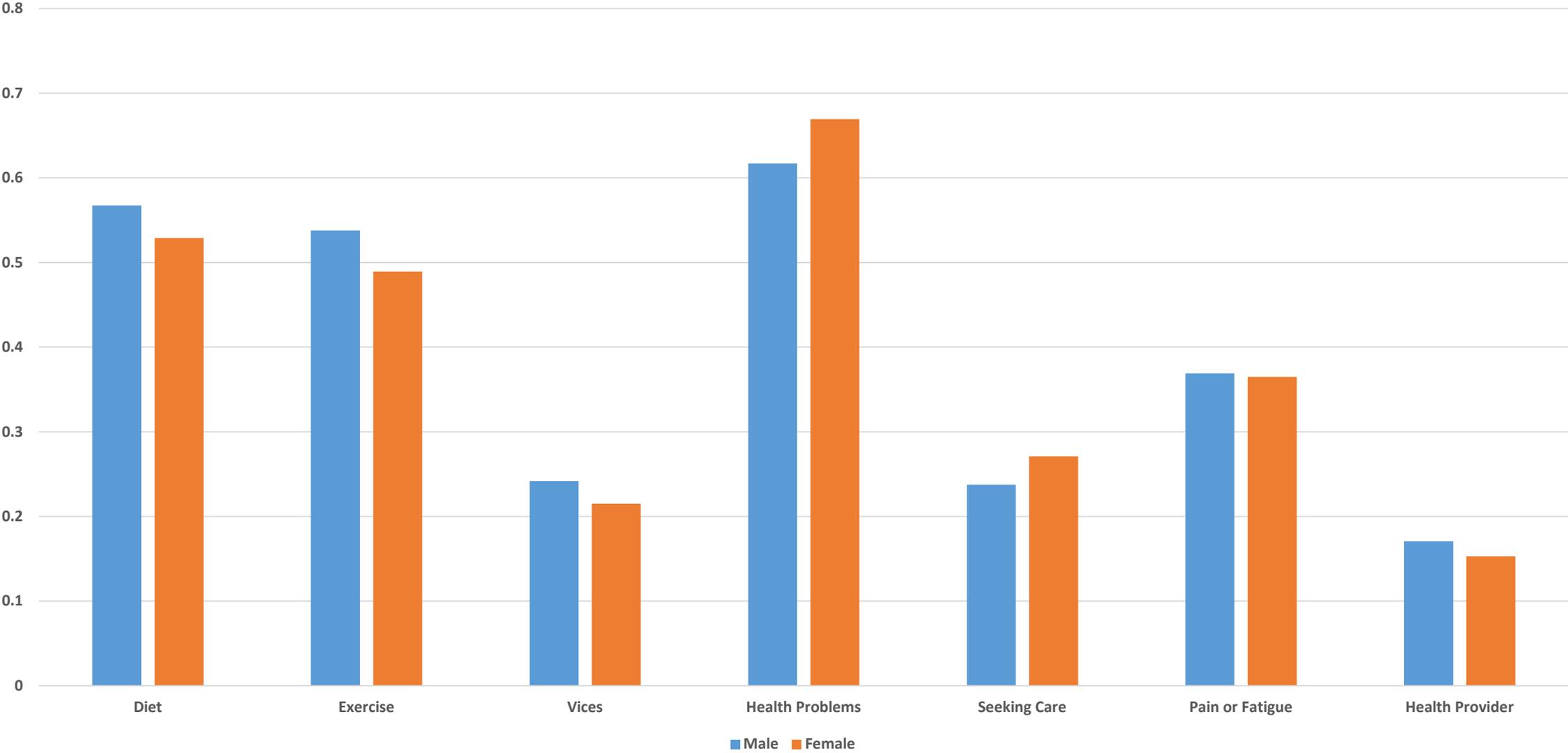
When you answered the previous question about your health, what did you think of? (Select all that apply)

1. My diet and nutrition
2. My exercise habits
3. My smoking or drinking habits
4. My health problems or conditions
5. The amount of times I seek health care
6. The amount of pain or fatigue that I have
7. My conversations with my doctor

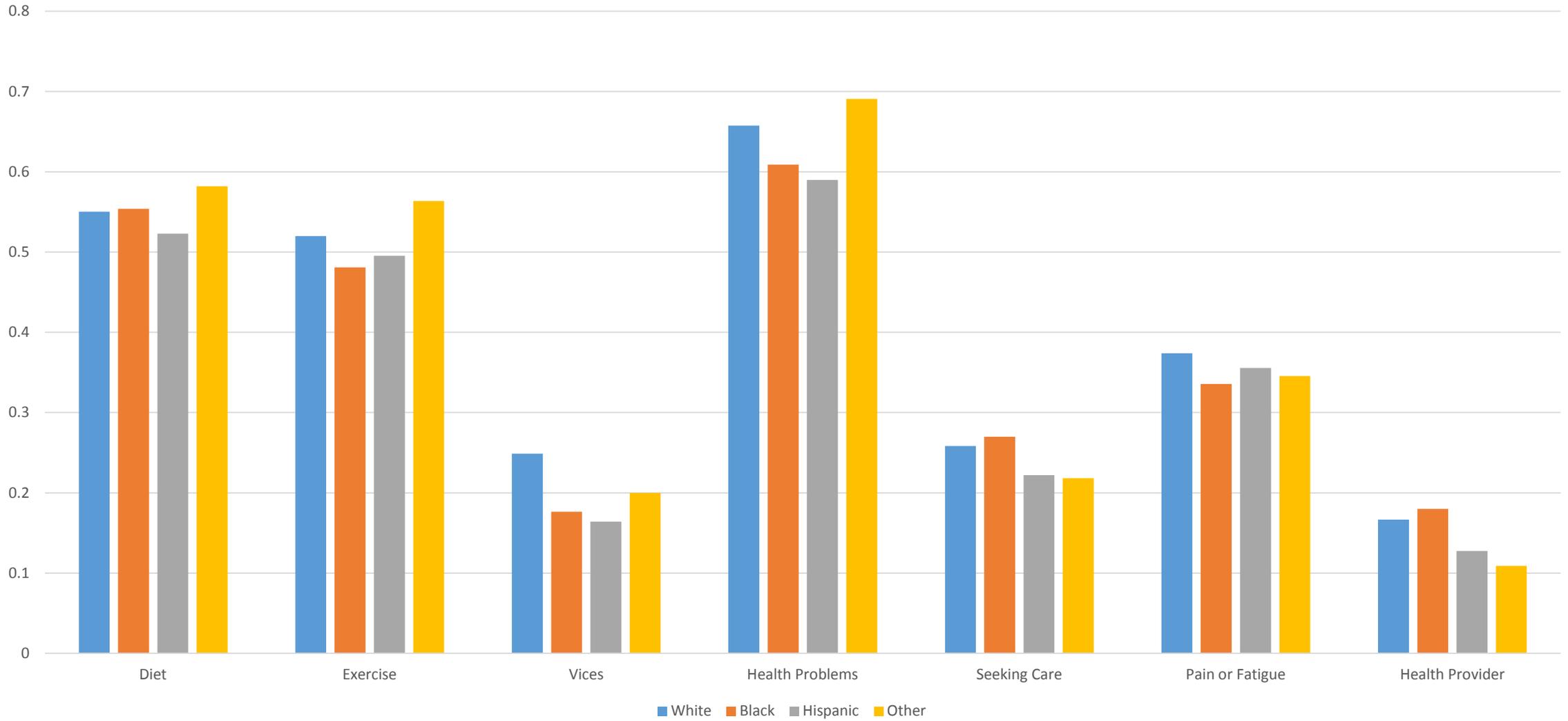
Proportion Indicating Each General Health Pattern of Interpretation



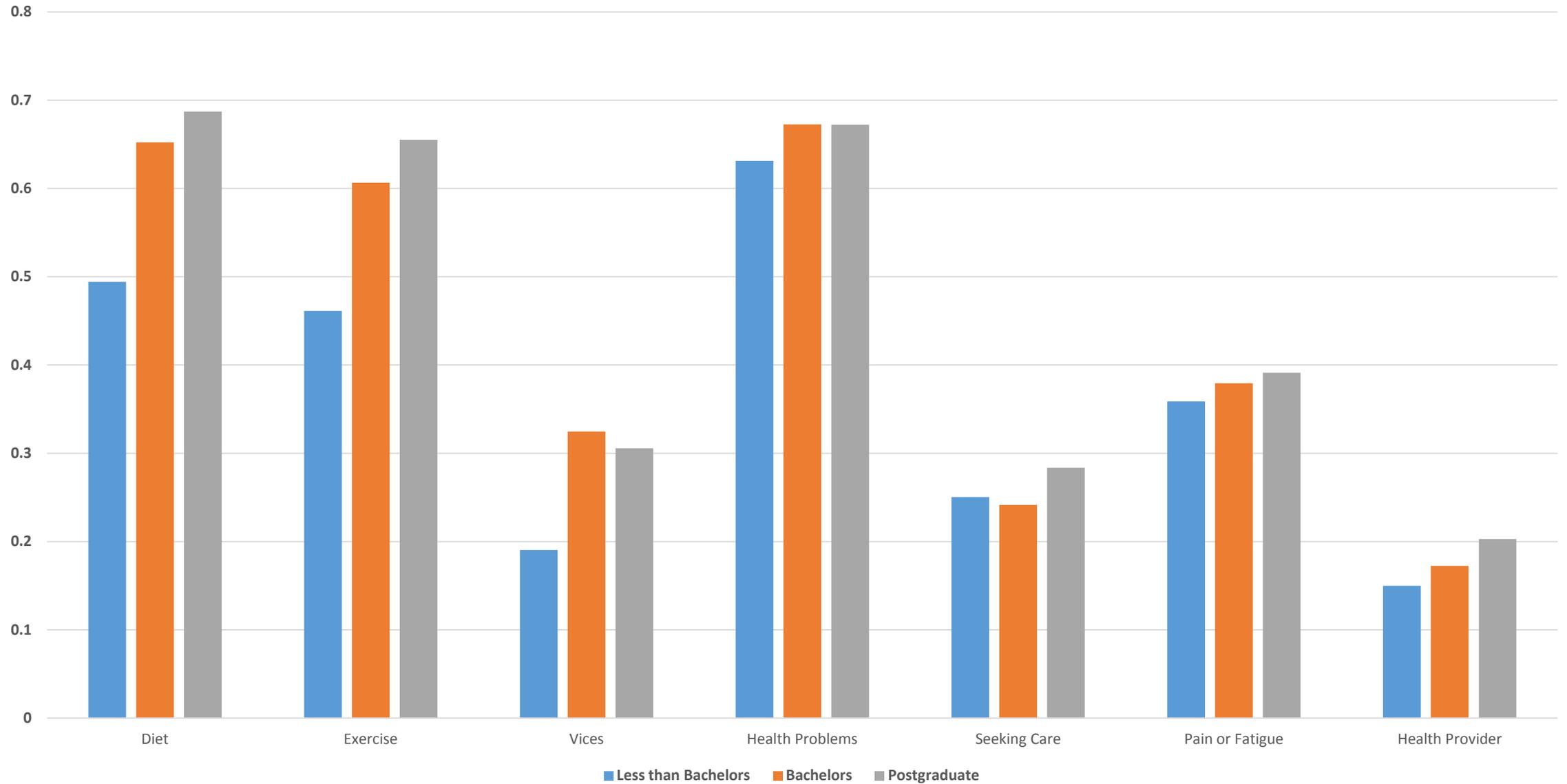
# General Health Patterns by Gender



## General Health Patterns by Race



## General Health Patterns by Educational Attainment



# Comparing the underlying constructs of two similarly-worded questions

## ■ Physical Activity questions

- Current NHIS asks about “Vigorous” and “Moderate” physical activity, as well as P.A. designed to strengthen muscles, using 3 separate questions.

VIG: How often do you do vigorous leisure-time physical activities for at least 10 minutes that cause heavy sweating or large increase in breathing or heart rate?

MOD: How often do you do light or moderate leisure time physical activities for at least 10 minutes that cause only light sweating or a slight to moderate increase in breathing or heart rate?

- One potential aspect of the NHIS redesign is to ask only one question for all types of physical activity.

NEW: In the past week, on how many days have you done a total of 30 minutes or more of physical activity, which was enough to raise your breathing rate? This may include sports, exercise, and brisk walking or cycling for recreation or to get to and from places, but should not include housework or physical activity that may be part of your job

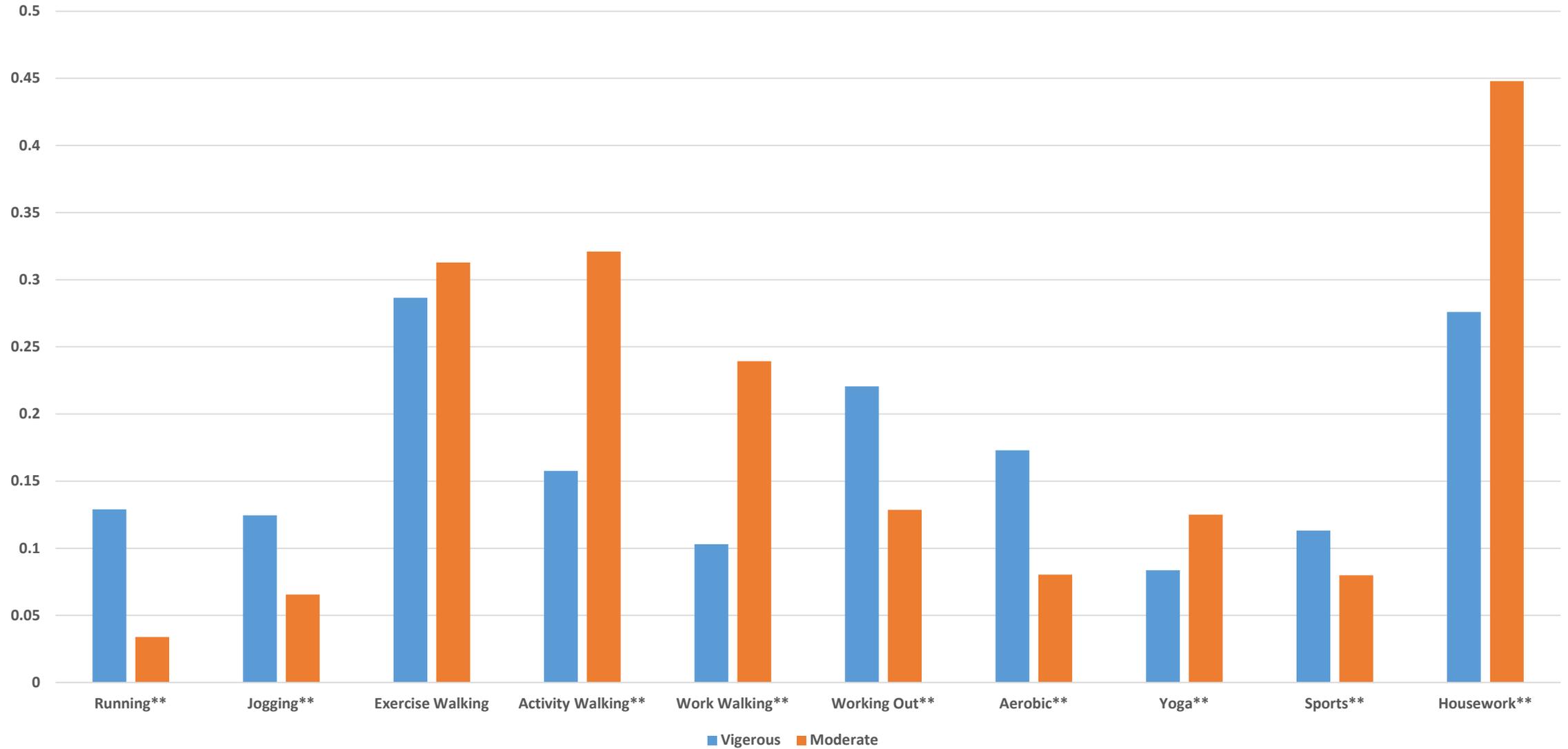
# Physical Activity Probe

- **Followed new, vigorous, and moderate questions:**

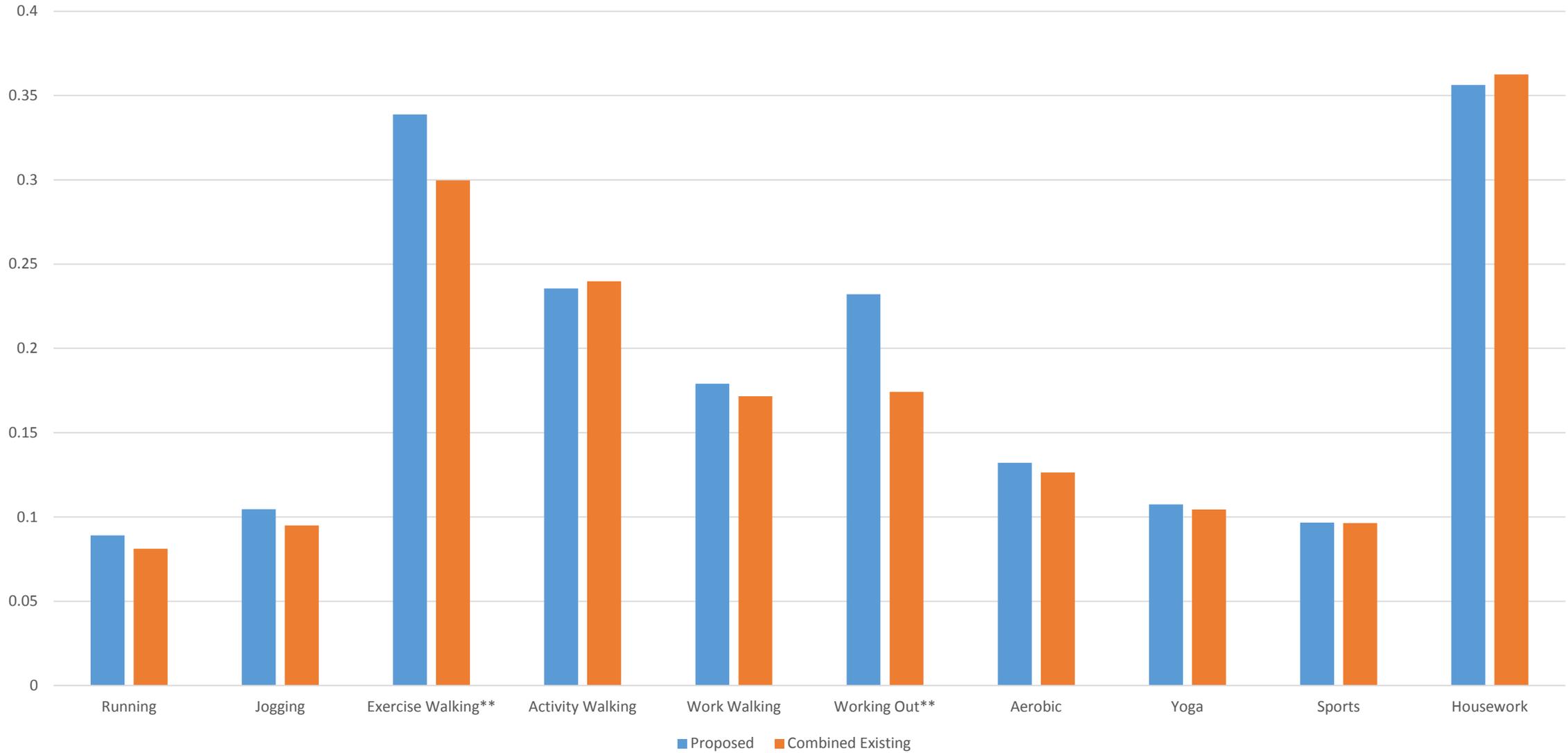
Which of the following types of physical activity, if any, did you include when you answered the previous question?  
(Select all that apply)

1. Running
2. Jogging
3. Walking or hiking for exercise
4. Walking to or from an activity
5. Walking at work
6. Working out with exercise equipment
7. Cycling, swimming, or other aerobic exercise
8. Yoga or Stretching
9. Playing sports
10. Housework or yardwork

## Constructs Captured by Existing NHIS Physical Activity Questions



Constructs Captured by Proposed and Existing NHIS Physical Activity Questions



# **Goal Two: Feasibility of Supplementing Traditional Surveys with Web Panels**

# Goal 2 Research to Date

- The basic scheme/strategy we are using is to divide NHIS questions into two categories:

(a) *core* – typically the more general questions, e.g.

Would you say your health... is excellent, ..... , or poor

... have you ever been told by a doctor...you have diabetes....

Have you smoked at least 100 cigarettes in your entire life

How often do you now smoke cigarettes.....

- (b) *detail* – specific questions that typically follow up the general questions, e.g.

How old were you when a doctor....told you that you had diabetes....

Are you now taking insulin?

Are you now taking diabetic pills to lower your blood sugar?...

**(dibpill)**

- If our experiment verifies feasibility, then we envision NHIS ultimately carried out in two parts:
  1. In person portion, containing Core Questions
  2. Web portion, containing Core and Detail Questions

- The core variables (responses to core questions) will serve as predictors/co-variates/ $x$ 's.
- The question is whether, with suitable estimation methods, we can predict well what the responses  $y$  to the detail questions on the NHIS would be, given:
  - (i) responses  $x$  to Core Questions in Web and NHIS
  - (ii) responses  $y$  to detail questions just on the Web.

# Main experiment (still in planning stage)

- (1) Get estimates for detail questions, using a variety of methods applied to Round 1 data and responses to Core Questions in Fall 2015 NHIS
- (2) Compare results to estimates derived directly from responses to Detail Questions on Fall 2015 NHIS
- (3) Assess which methods work and work best
- (4) Check conclusions using Round 2, comparing to Spring 2016 NHIS

# The Questions:

1. What estimation methods?
2. How evaluate the various methods (what metrics)?

# Methods

## I. *Univariate* (e.g. basing estimates to **dibpill** on just web responses to **dibpill**)

Unweighted; weighted using variety of approaches: Gallup supplied weights, post-stratification, propensity score, other?

Propensity score estimated parametrically (logistic regression) or non-parametrically (generalized boosted regression models)

## II. *Incorporating co-variates* (x's from Core Questions plus demographic variables)

Parametric, non-parametric – weighted and unweighted

For each Method, we want to have a concomitant **Variance Estimation** Methodology or other measure of accuracy.

(We will not be able to produce official statistics without being able to offer estimates of precision or accuracy)

Our evaluation will include an appraisal of the variance estimation.

## Evaluation (still under discussion)

Suppose there are  $n_d$  detail questions

Let  $\hat{y}_j$  be NHIS estimate for  $j$ th question (a kind of target)

$\tilde{y}_{\nu j}$  be Web estimate for  $j$ th question, using the  $\nu$ th method

Some possible methods of evaluation:

eyeballing

for example graphing  $\tilde{y}_{\nu j}$  versus  $\hat{y}_j$  for different methods  $\nu$

average of biases or absolute biases across  $\nu$ 's  $n_d^{-1} \sum_j |\tilde{y}_{\nu j} - \hat{y}_j|$

(cf. PewResearchCenter Report, May 2 2016)

average of relative absolute biases  $n_d^{-1} \sum_j |\tilde{y}_{\nu j} / \hat{y}_j - 1|$

average of comparative absolute biases

$$n_d^{-1} \sum_j \frac{|\tilde{y}_{vj} - \hat{y}_j|}{\min_{v'} |\tilde{y}_{v'j} - \hat{y}_j|}$$

weighted versions of above....

Distribution of  $\frac{\tilde{y}_{vj} - \hat{y}_j}{\sqrt{\text{var}_v(\tilde{y}_{vj}) + \text{var}_v(\hat{y}_j)}}$

Versions tailored to estimating proportions....

# Preliminary Experiments using just NHIS data

- Extraction of Feb – June 2014 (“NHIS Target”) and December 2014 (“Web”)
- Two versions of “Web”:
  - (a) limited to those whose answers to “use of Internet” was positive
  - (b) all of December

- Purpose of Preliminary Experiments:
  - Practice application of methods
  - See how well methods do without confounding factor of panel selection and possibly mode effect
  - Gives background against which to judge results in Main Experiment
  - Get clarification on different sorts of models, which variables to use, etc.

Estimation from Web Panel Surveys: Based on Simulated Data from NHIS

Table Rand_S1. Key Estimates from NHIS and Web panel Surveys							
Response Item	a	b	c	d	e	f	g
	NHIS: Unweighted(n)	NHIS: Weighted by final weights for sample adults (wtfa_sa)	Web: Unweighted(n)	Web: Weighted by post- stratified weights. See Note 1	Web: Weighted by propensity weights (Via unweighted logistic regression). See Note 2)	Web: Weighted by propensity weights (Via weighted logistic regression). See Note 3)	Web: Weighted by propensity weights (Via weighted logistic regression). See Note 4)
% Medicaid coverage (hikindnd)	10.12 (15,282)	8.70	8.69 (1,980)	8.42	8.50	8.54	9.09
% Ever been told that you have diabetes (dibev) - "borderline" regarded as "DK/Refused"	10.62 (15,043)	9.29	7.07 (1,965)	7.48	7.96	7.63	7.84
% Now taking diabetic pill (dibpill)	51.56 (2,401)	52.35	50.00 (226)	51.13	51.62	51.06	51.49
% Ever had prediabetes or other symptoms (dibpre1)	5.86 (13,719)	5.35	4.71 (1,846)	4.89	5.08	4.92	4.89
Age (mean) first diagnosed w/diabetes (dibage)	48.09 (1,564)	47.64	46.41 (138)	46.68	47.14	47.00	46.82
Notes: 1) Post-stratified by age (18-34; 35-54; 55+), sex, race/ethnicity (Hispanic, NH White, NH Black, NH Other). The control totals were estimated using NHIS final weights for sample adults; 2) Unweighted logistic regression was fitted with the same set of explanatory variables used in 1); 3) Weighted logistic regression was fitted with the same set of explanatory variables used in 1) with post-stratified Web weights (to "internet" population) and NHIS final weights for sample adults. Distribution of "internet" population was estimated from NHIS; 4) The same method as in 3) with additional two variables - poverty indicator (Under poverty ratio vs. Other) and 4-category education..							

- Other preliminary investigation also being carried out using actual Gallup data conjoined with 2014 data
- Some other questions/issues
  1. Which should be core questions/ which detail
  2. How handle NA's some of which arise from skip patterns, some from various forms of non-response
  3. Some of skip patterns did not exactly match between original NHIS and Web
  4. Overarching Question: characterizing/reconciling the populations targeted by the NHIS and Web survey