Formative and User Study Materials:

A. **Script for Semi-Structured Interviews**
The script contains a general structure and prompts for the formative interviews. The interviews were otherwise organic and followed the conversation of the participant, with the script guiding the topic of conversation.

B. **Example Visualizations for Semi-Structured Interviews**
The visualizations show the Vega specification, output visualization, and corresponding dataflow graph. The dataflow graph is a representation of Vega's system internals and shows the process with which the output Vega visualization is constructed. The dataflow graph was a prototype of a possible visualization for the system.

C. **Reference Sheet for User Study**
A printed version of the reference sheet was given to each participant in the user study for Vega's visual debugging techniques. It was used as a reference for the techniques available during the tasks and during the follow-up questions to help remind users of the names for each technique.

Interview Questions

The goal of this interview is to gather information on the development and debugging process of Vega users.

1. **Walk me through the process you use when working with Vega.**
2. **What version of Vega are you currently using?**
3. *(if not answered above)* **What is your primary method for producing Vega specifications?** To clarify, methods could include: (1) by hand, (2) modifications to existing specifications, (3) programmatically generated by your own code, (4) programmatically generated by another system
4. **What was the last (or most troublesome) error you encountered when generating a Vega specification?**
   a. *(if not answered above)* What was your approach towards resolving this problem?
   b. *(if not answered above)* What information was most important in identifying the problem?
   c. *(if not answered above)* What tools did you use to facilitate the debugging process?
5. **In what ways do you think the debugging process could have been facilitated?**

6. **Are there any other errors that you found particularly difficult to debug?** *Repeat Q5 and Q6.*

Up to this point, we've mainly discussed your current strategies when developing and debugging Vega. At this point, I would like to dive a little deeper into what strategies or techniques could be added to facilitate the development process.

7. **In general, what additional functionality or information could be added to facilitate your development process?**
8. One debugging strategy is to use the console to inspect the underlying execution structure and scenegraph of Vega. **How often do you use the console for debugging Vega?**
   a. *(if not discussed above)* To what extent do you examine the underlying execution structure of Vega?
   b. *(if not aware)* Do you think this structure is something you could/will use in the future for debugging Vega specifications? If so, how? If not, why not?

At this point, show a pre-selected specification, visualization, and DFG.

The JavaScript console allows the user to access a lot of potentially useful information, but may require more domain knowledge or an intricate understanding of the Vega internals. One way

A. Script for Semi-Structured Interviews

to reduce the difficulty associated with inspecting the textual structure is to provide a visualization of the underlying structure.

This example shows the spec, visualization, and underlying dataflow graph **(See Part B)**.

9. **What is your initial impression of the dataflow graph?** In particular, what do you notice? What are you confused by?

Based on the participants impressions and understanding of the dataflow graph, provide more context and discuss. Examples of what to point out could include:

Looking at this graph, we can notice patterns in the structure that correspond to patterns in the spec, data, and visualization… (Point out things like facet, bars+text, potentially show colored version below at a later point).

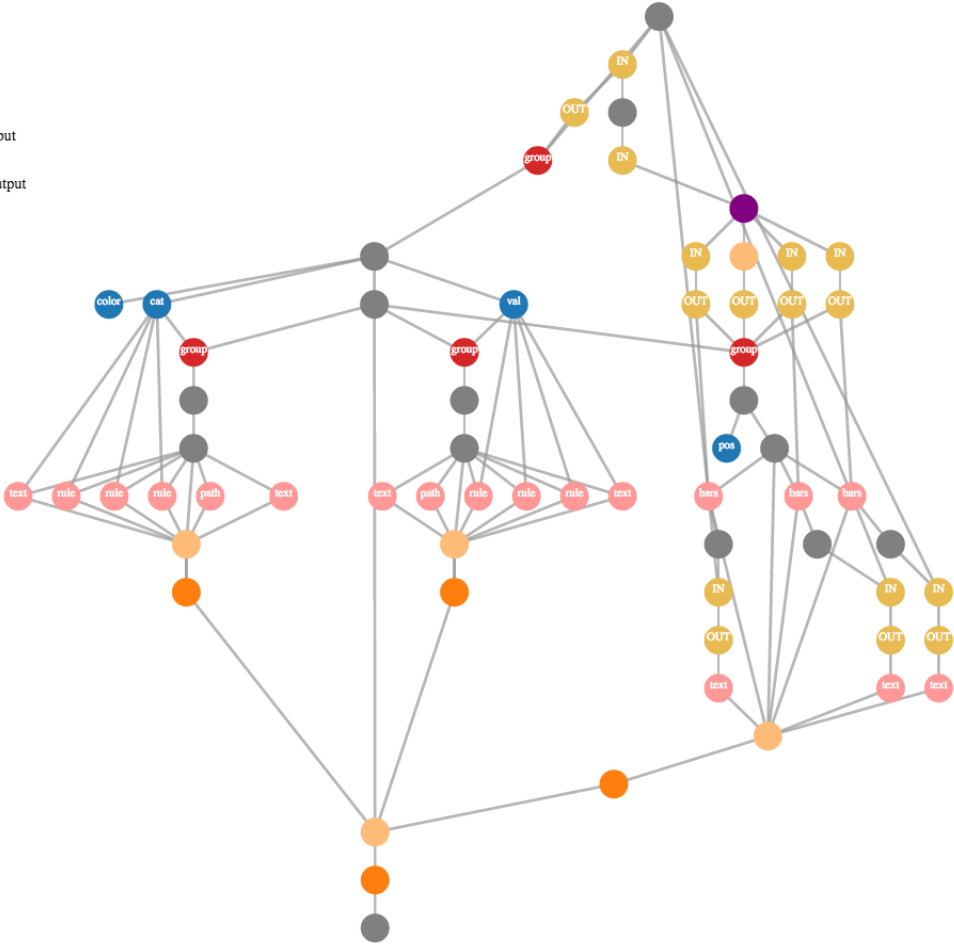10. **Do you think that having this structure available would be useful for the development process?**
    a. *(if yes)* What parts of this representation do you think would be most useful? What sorts of interactions or information do you think would be relevant or helpful?
    b. *(if no)* Why do you think this representation would not be useful? Is there any information within this structure that would be useful to surface in another way?
11. **What do you think could be added or changed to improve this visual representation?**

Those are all the questions I have. Thank you for taking the time to discuss these problems/topics with me. Please feel free to reach out if you have any additional thoughts on the topic or uncover other debugging scenarios that could be facilitated.
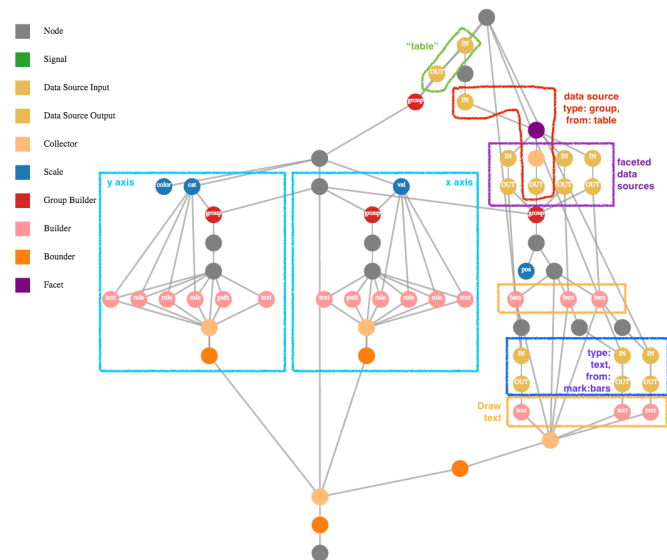
B. Example Visualizations for Semi-Structured Interviews

```json
{ "width": 300,
  "height": 240,
  "data": [{ "name": "table",
     "values": [
        {"category":"A", "position":0, "value":0.1},
        {"category":"A", "position":1, "value":0.6},
        {"category":"A", "position":2, "value":0.9},
        {"category":"A", "position":3, "value":0.4},
        {"category":"B", "position":0, "value":0.7},
        {"category":"B", "position":1, "value":0.2},
        {"category":"B", "position":2, "value":1.1},
        {"category":"B", "position":3, "value":0.8},
        {"category":"C", "position":0, "value":0.6},
        {"category":"C", "position":1, "value":0.1},
        {"category":"C", "position":2, "value":0.2},
        {"category":"C", "position":3, "value":0.7}]
  }],
  "scales": [{ "name": "cat", "type": "ordinal",
     "range": "height", "padding": 0.2,
     "domain": {"data": "table", "field": "category"}
  },
  { "name": "val", "range": "width",
     "round": true, "nice": true,
     "domain": {"data": "table", "field": "value"}
  },
  { "name": "color", "type": "ordinal", "range": "category20"
     "domain": {"data": "table", "field": "position"}
  }],
  "axes": [
  {"type": "y", "scale": "cat", "tickSize": 0, "tickPadding": 8},
  {"type": "x", "scale": "val"}
  ],
  "marks": [{ "type": "group",
     "from": {
        "data": "table",
        "transform": [{"type":"facet", "keys":["category"]}]
     },
     "properties": { "enter": {
        "y": {"scale": "cat", "field": "key"},
        "height": {"scale": "cat", "band": true}}
     },
     "scales": [{
        "name": "pos", "type": "ordinal", "range": "height",
        "domain": {"field": "position"}
     }],
     "marks": [{ "name": "bars", "type": "rect",
        "properties": { "enter": {
           "y": {"scale": "pos", "field": "position"},
           "height": {"scale": "pos", "band": true},
           "x": {"scale": "val", "field": "value"},
           "x2": {"scale": "val", "value": 0},
           "fill": {"scale": "color", "field": "position"}}
     }},
     { "type": "text", "from": {"mark": "bars"},
        "properties": { "enter": {
           "x": {"field": "x2", "offset": -5},
           "y": {"field": "y"},
           "dy": {"field": "height", "mult": 0.5},
           "fill": {"value": "white"},
           "align": {"value": "right"},
           "baseline": {"value": "middle"},
           "text": {"field": "datum.value"}
     }}}
]}]}
```

B. Example Visualizations for Semi-Structured Interviews
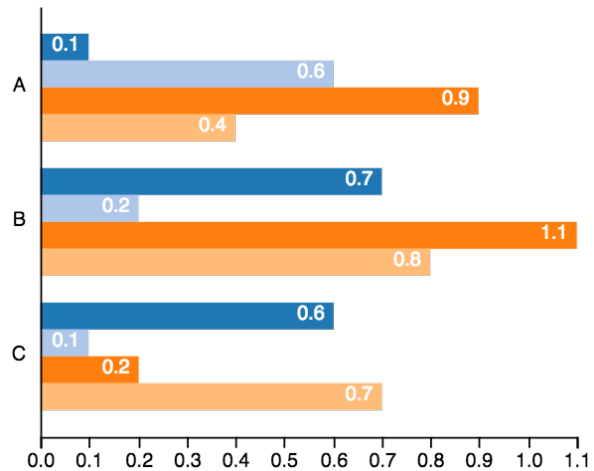
B. Example Visualizations for Semi-Structured Interviews

```json
{ "width": 300,
  "height": 240,
  "data": [{ "name": "table",
    "values": [
      {"category":"A", "position":0, "value":0.1},
      {"category":"A", "position":1, "value":0.6},
      {"category":"A", "position":2, "value":0.9},
      {"category":"A", "position":3, "value":0.4},
      {"category":"B", "position":0, "value":0.7},
      {"category":"B", "position":1, "value":0.2},
      {"category":"B", "position":2, "value":1.1},
      {"category":"B", "position":3, "value":0.8},
      {"category":"C", "position":0, "value":0.6},
      {"category":"C", "position":1, "value":0.1},
      {"category":"C", "position":2, "value":0.2},
      {"category":"C", "position":3, "value":0.7}]
  }],
  "scales": [{ "name": "cat", "type": "ordinal",
    "range": "height", "padding": 0.2,
    "domain": {"data": "table", "field": "category"}
   },
   { "name": "val", "range": "width",
     "round": true, "nice": true,
     "domain": {"data": "table", "field": "value"}
   },
   { "name": "color", "type": "ordinal", "range": "category20"
     "domain": {"data": "table", "field": "position"}
  }],
  "axes": [
   {"type": "y", "scale": "cat", "tickSize": 0, "tickPadding": 8},
   {"type": "x", "scale": "val"}
  ],
  "marks": [{ "type": "group",
    "from": {
      "data": "table",
      "transform": [{"type":"facet", "keys":["category"]}]
    },
    "properties": { "enter": {
      "y": {"scale": "cat", "field": "key"},
      "height": {"scale": "cat", "band": true}}
    },
    "scales": [{ "name": "pos", "type": "ordinal", "range": "height",
      "domain": {"field": "position"}
    }],
    "marks": [{ "name": "bars", "type": "rect",
      "properties": { "enter": {
        "y": {"scale": "pos", "field": "position"},
        "height": {"scale": "pos", "band": true},
        "x": {"scale": "val", "field": "value"},
        "x2": {"scale": "val", "value": 0},
        "fill": {"scale": "color", "field": "position"}}
    }},
    { "type": "text", "from": {"mark": "bars"},
      "properties": { "enter": {
        "x": {"field": "x2", "offset": -5},
        "y": {"field": "y"},
        "dy": {"field": "height", "mult": 0.5},
        "fill": {"value": "white"},
        "align": {"value": "right"},
        "baseline": {"value": "middle"},
        "text": {"field": "datum.value"}
      }}}
  ]}]}
```
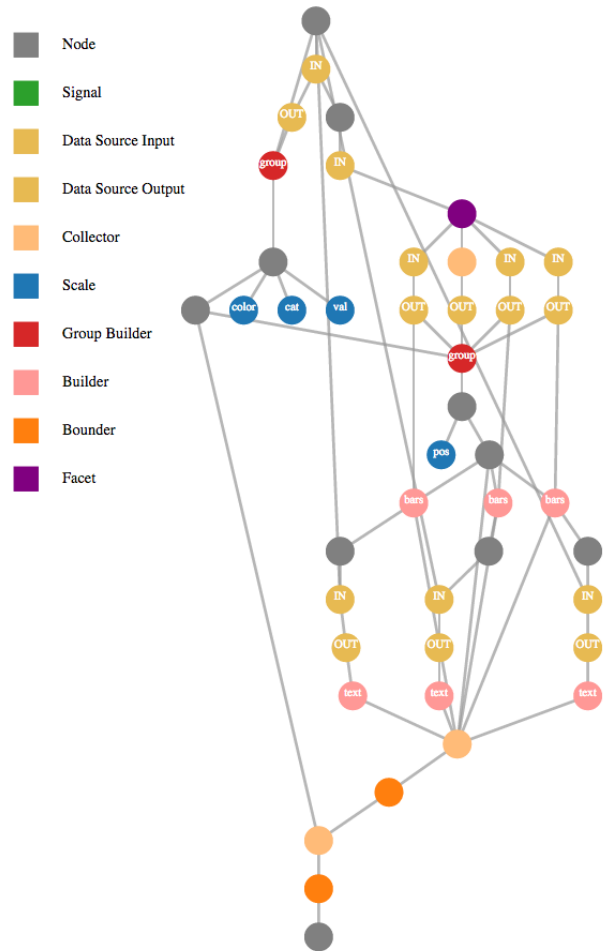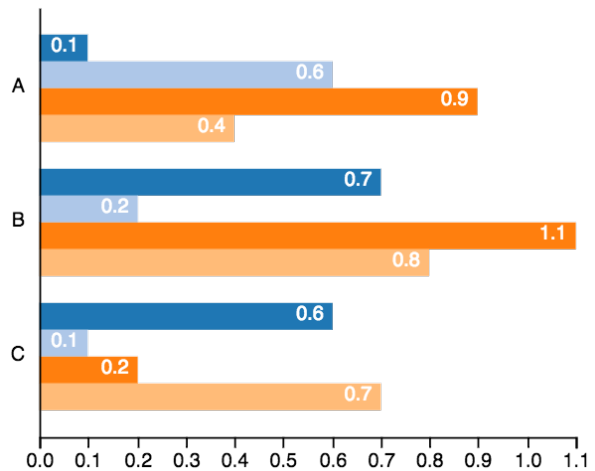
B. Example Visualizations for Semi-Structured Interviews

```json
{ "width": 300,
  "height": 240,
  "data": [{ "name": "table",
    "values": [
      {"category":"A", "position":0, "value":0.1},
      {"category":"A", "position":1, "value":0.6},
      {"category":"A", "position":2, "value":0.9},
      {"category":"A", "position":3, "value":0.4},
      {"category":"B", "position":0, "value":0.7},
      {"category":"B", "position":1, "value":0.2},
      {"category":"B", "position":2, "value":1.1},
      {"category":"B", "position":3, "value":0.8},
      {"category":"C", "position":0, "value":0.6},
      {"category":"C", "position":1, "value":0.1},
      {"category":"C", "position":2, "value":0.2},
      {"category":"C", "position":3, "value":0.7}]
  }],
  "scales": [{ "name": "cat", "type": "ordinal",
    "range": "height", "padding": 0.2,
    "domain": {"data": "table", "field": "category"}
  },
  { "name": "val", "range": "width",
    "round": true, "nice": true,
    "domain": {"data": "table", "field": "value"}
  },
  { "name": "color", "type": "ordinal", "range": "category20"
    "domain": {"data": "table", "field": "position"}
  }],
  "axes": [
  {"type": "y", "scale": "cat", "tickSize": 0, "tickPadding": 8},
  {"type": "x", "scale": "val"}
  ],
  "marks": [{ "type": "group",
    "from": {
      "data": "table",
      "transform": [{"type":"facet", "keys":["category"]}]
    },
    "properties": { "enter": {
      "y": {"scale": "cat", "field": "key"},
      "height": {"scale": "cat", "band": true}}
    },
    "scales": [{
      "name": "pos", "type": "ordinal", "range": "height",
      "domain": {"field": "position"}
    }],
    "marks": [{ "name": "bars", "type": "rect",
      "properties": { "enter": {
        "y": {"scale": "pos", "field": "position"},
        "height": {"scale": "pos", "band": true},
        "x": {"scale": "val", "field": "value"},
        "x2": {"scale": "val", "value": 0},
        "fill": {"scale": "color", "field": "position"}}
    }},
    { "type": "text", "from": {"mark": "bars"},
      "properties": { "enter": {
        "x": {"field": "x2", "offset": -5},
        "y": {"field": "y"},
        "dy": {"field": "height", "mult": 0.5},
        "fill": {"value": "white"},
        "align": {"value": "right"},
        "baseline": {"value": "middle"},
        "text": {"field": "datum.value"}
      }}}
  ]}]}
```

Replay Indicator:

🎯 *record: interact with the vis. which updates timeline*

🔄 *replay: display past state and **tooltip***

Dependencies:
*highlights the signal & value on which used to define the signal at **mouseover (dark gray)***

Change Summary: *(only with data tab) the line charts show how the distribution of values in each property (id, _id) changes over time.*
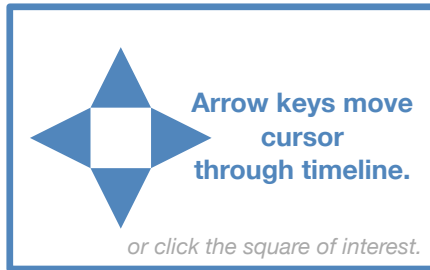
iris	selectedPoints
id	_id
42	338
132	339
118	340
110	352
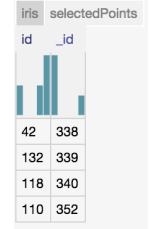
Overview:
bar height shows # signals changed in pulse

Timeline:
squares show user defined signals over time

Vega Timeline Data 🔄   shiftclick_select ▾   Format   Parse

width		200
height		200
clickedPoint		110
shift 🔗		true
tooltip		{"x":134,"y":35,"_i...
sepalWidth		3.6080000000...

Tooltip: *(only during replay) information about the underlying visualization state*

```
pixel    scale    data
115  →  x⁻¹  ⇒  3.38
109  →  y⁻¹  ⇒  3.73
```

Specification:
user defined code for visualization

```
Vega Timeline Data
1  {
2
3      "width": 200,
4      "height": 200,
5
6      "signals": [
7          {"name": "width", "init": 200},
8          {"name": "height", "init": 200},
9          {
10             "name": "clickedPoint",
11             "init": 0,
12             "verbose": true,
13             "streams": [{"type": "click", "expr": "datum._id"}]
14         },
15         {
16             "name": "shift",
17             "init": false,
18             "verbose": true,
19             "streams": [{"type": "click", "expr": "event.shiftKey"}]
20         },
21         {
22             "name": "tooltip",
23             "init": {},
24             "streams": [
25                 {"type": "symbol:mouseover", "expr": "{'x': event.x(), 'y': event.y(), '_id'
26                 {"type": "symbol:mouseout", "expr": "{}"}
27             ]
28         },
29         {
```
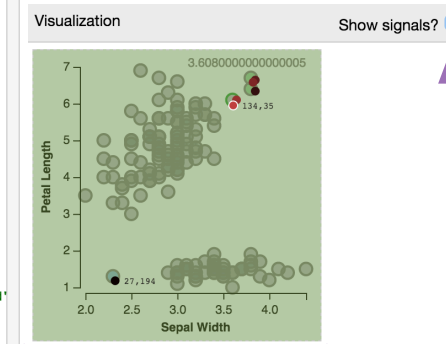
Arrow keys move cursor through timeline.

or click the square of interest.

Visualization — 3.6080000000000005 — Petal Length / Sepal Width — Show signals?

Visualization — 3.6080000000000005 — Petal Length / Sepal Width — Show signals? ✓ — 134,35 — 27,194

Data Table:
selecting the data tab shows the data table.

tabs at the top exist for each user-defined data set in the spec.

histograms show the distribution of value in at the current time.

iris	selectedPoints					
sepalLength	sepalWidth	petalLength	petalWidth	species	_id	_prev
5.1	3.5	1.4	0.2	"setosa"	1	{"_id":1,"sepalLength":5.1,"...
4.9	3	1.4	0.2	"setosa"	2	{"_id":2,"sepalLength":4.9,...
4.7	3.2	1.3	0.2	"setosa"	3	{"_id":3,"sepalLength":4.7,"...
4.6	3.1	1.5	0.2	"setosa"	4	{"_id":4,"sepalLength":4.6,"...
5	3.6	1.4	0.2	"setosa"	5	{"_id":5,"sepalLength":5,"s...
5.4	3.9	1.7	0.4	"setosa"	6	{"_id":6,"sepalLength":5.4,"...
4.6	3.4	1.4	0.3	"setosa"	7	{"_id":7,"sepalLength":4.6,"...
5	3.4	1.5	0.2	"setosa"	8	{"_id":8,"sepalLength":5,"s...
4.4	2.9	1.4	0.2	"setosa"	9	{"_id":9,"sepalLength":4.4,"...
4.9	3.1	1.5	0.1	"setosa"	10	{"_id":10,"sepalLength":4.9...

Click for more data.

Cursor:
the current point in the timeline. the most recent value for a given signal is shown in light green and the value is displayed in the rightmost column.

Annotations:
Signals corresponding to a point are drawn on the visualization.

The current point is bordered in white and colored based on:
dark (past) > red (current) > white (future)